

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

ON protein - protein search, using sw model

Run on: May 1, 2003, 22:05:56 ; Search time 29 Seconds
(without alignments)
231.325 Million cell updates/sec

Title: US-09-821-255-2

Perfect score: 1211

Sequence: 1 ILGREGREAHARPMASVOL.....KPGIYTRVASYAAWIDSVLA 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

al number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTDATA_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score of at least the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1205	99.5	238	6	5223425-5
2	1201	99.5	250	6	5223425-4
3	1201	99.5	253	6	5223425-8
4	1198	98.8	228	4	US-08-944-483-44
5	1051.5	86.8	222	6	5223425-6
6	844.5	69.7	259	6	5223425-5
7	844	69.7	260	6	5223425-10
8	475.5	37.9	232	4	US-08-944-483-45
9	454.5	37.9	232	4	US-08-944-483-46
10	452.5	37.4	232	1	US-07-990-301A-4
11	445	36.7	258	1	US-09-008-271A-7
12	440.5	36.4	260	4	US-08-361-395-1
13	433.5	35.8	244	1	US-08-944-483-37
14	433.5	35.8	260	3	US-09-070-526-2
15	421	34.8	288	4	US-09-386-642-13
16	419.5	34.6	260	3	US-09-025-059-3
17	419.5	34.6	260	3	US-09-025-059-3
18	415	34.3	237	3	US-08-768-859A-16
19	415	34.3	237	3	US-08-768-859A-16
20	415	34.3	237	3	US-08-622-046B-12
21	415	34.3	237	4	US-08-944-483-37
22	415	34.3	237	4	US-09-100-264-1
23	415	34.3	238	3	US-08-768-859A-8
24	415	34.3	238	3	US-08-767-820A-8
25	415	34.3	243	3	US-08-767-820A-10
26	415	34.3	244	3	US-08-622-046B-10
27	415	34.3	244	3	US-08-622-046B-16

28 415 34.3 244 4 US-09-100-264-5
29 415 34.3 261 3 US-08-768-859A-6
30 415 34.3 261 3 US-08-767-820A-6
31 415 34.3 261 3 US-08-767-820A-6
32 415 34.3 261 3 US-08-767-820A-6
33 415 34.3 261 4 US-08-983-073D-7
34 415 34.3 261 4 US-08-983-073D-7
35 415 34.3 261 4 US-08-983-073D-7
36 415 34.3 261 4 US-08-983-073D-7
37 415 34.3 261 4 US-08-983-073D-7
38 415 34.3 261 4 US-08-983-073D-7
39 415 34.3 261 4 US-08-983-073D-7
40 415 34.3 261 4 US-08-983-073D-7
41 413.5 34.1 237 3 US-08-622-046B-7
42 413.5 34.1 237 3 US-08-622-046B-7
43 413.5 34.1 237 3 US-08-622-046B-7
44 413.5 34.1 237 3 US-08-622-046B-7
45 413.5 34.1 237 3 US-08-622-046B-7

ALIGNMENTS

RESULT 1
5223425-5
; PATENT NO. 5223425
; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,
; BARRY M.; WHITE, TYLER R.
; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT
; D ACTIVITY
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA: /07/277,963
; FILING DATE: 30-NOV-1988
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 34,203
; FILING DATE: 02-APR-1987
; SEQ ID NO: 5
; LENGTH: 238
5223425-5

Query Match 99.5%; Score 1205; DB 6; Length 238;
Best Local Similarity 99.6%; Pred. No. 1.3e-120;
Matches 227; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILGREGREAHARPMASVOLGREGVAVAPVLSNAHCLDEADADKSVLLGAHSL 60
DB 11 ILGREGREAHARPMASVOLGREGVAVAPVLSNAHCLDEADADKSVLLGAHSL 70
OY 61 SOPEPSKRLVDLRAVPHPSQPTDTHDILLQSLSEKATIGPVPRLPMQVDRVAPG 120
DB 71 SOPEPSKRLVDLRAVPHPSQPTDTHDILLQSLSEKATIGPVPRLPMQVDRVAPG 130
OY 131 TLCDVAGMIVNHACRPSDLSQHVLLVPLVDLRATCNRRTHDGAITERLMCAESNRDSC 180
DB 131 TLCDVAGMIVNHACRPSDLSQHVLLVPLVDLRATCNRRTHDGAITERLMCAESNRDSC 190
OY 181 GDGGGVLGCGVLEGVWTSRVCGRNKKPGIYTRVASYAAWIDSVLA 228
DB 191 GDGGGVLGCGVLEGVWTSRVCGRNKKPGIYTRVASYAAWIDSVLA 238

RESULT 2
5223425-8
; PATENT NO. 5223425
; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,
; BARRY M.; WHITE, TYLER R.
; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT
; D ACTIVITY
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA: /08/07277,963
; FILING DATE: 30-NOV-1988

RESULT 4

RESULT 7
5223425-10
: TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT
: D ACTIVITY
: PATENT NO. 5223425
: INVENTOR: JEFFREY S.: SPIEGELMAN, BRUCE M.: ROSEN,
: BARRY M.: WHITE, TYLER H.
: TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT
: D ACTIVITY
: NUMBER OF SEQUENCES: 19
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/277,963
: FILING DATE: 30-NOV-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 34,203
: FILING DATE: 02-APR-1987
: SEQ ID NO:6
: LENGTH: 222
5223425-6

Query Match 86.7%; Score 1051.5; DB 6; Length 222;
Best Local Similarity 89.9%; Pred. No. 2.8e-104;
Matches 205; Conservative 4; Mismatches 12; Indels 7; Gaps 3;
QY 1 ILGREAHAARPYMASVOLNGAHLCGGVLAQWVLSAARHLEDAADGKQVOLLGAHSL 60
1 ILGREAHAARPYMASVOLNGAHLCGGVLAQWVLSAARHLEDAADGKQVOLLGAHSL 60
61 SQPSPEKLLTYLRAVPHDSQPTIDHLLQLSEKATLGPVAPRLPMQVDRDVP 119
61 POPEPXKXITLRAVPHDSQPTIDHLLQLSEKATLGPVAPRLPMQVDRDVP 120
DB 120 GTLCDVAGCIVNHAAGRPDSLOHVLPLVDORATCNRRTHGDCAITERLMAESNRDSC 179
121 GTLCDVAGCIVNHAAGRPDSLOHVLPLVDORATCNRRTHGDCAITERLMAESNRDSC 174
121 GTLCDVAGCIVNHAAGRPDSLOHVLPLVDORATCNRRTHGDCAITERLMAESNRDSC 174
QY 180 KDSGSGPLVCGGVLGVVTHSGSVCCGKPKGIVTVASVAAIDSVL 227
180 KDSGSGPLVCGGVLGVVTHSGSVCCGKPKGIVTVASVAAIDSVL 222
DB 175 KDSGSGPLVCGGVLGVVTHSGSVCCGKPKGIVTVASVAAIDSVL 222

RESULT 6
5223425-2
: TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT
: D ACTIVITY
: PATENT NO. 5223425
: INVENTOR: JEFFREY S.: SPIEGELMAN, BRUCE M.: ROSEN,
: BARRY M.: WHITE, TYLER H.
: TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT
: D ACTIVITY
: NUMBER OF SEQUENCES: 19
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/277,963
: FILING DATE: 30-NOV-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 34,203
: FILING DATE: 02-APR-1987
: ID NO:2
: LENGTH: 259
5223425-2

Query Match 69.7%; Score 844.5; DB 6; Length 259;
Best Local Similarity 66.2%; Pred. No. 4.5e-82;
Matches 151; Conservative 33; Mismatches 42; Indels 1; Gaps 1;
QY 1 ILGREAHAARPYMASVOLNGAHLCGGVLAQWVLSAARHLEDAADGKQVOLLGAH 59
1 ILGREAHAARPYMASVOLNGAHLCGGVLAQWVLSAARHLEDAADGKQVOLLGAH 59
26 ILGREAHAARPYMASVOLNGAHLCGGVLAQWVLSAARHLEDAADGKQVOLLGAH 85
60 LSQPSPEKLLTYLRAVPHDSQPTIDHLLQLSEKATLGPVAPRLPMQVDRDVP 119
60 LSQPSPEKLLTYLRAVPHDSQPTIDHLLQLSEKATLGPVAPRLPMQVDRDVP 119
DB 86 LSAPPEYKRWTDVQSVVPHGSPDSDLEDDLLFKLSQNASLGRHVRPLPYEDKEVEP 145
86 LSAPPEYKRWTDVQSVVPHGSPDSDLEDDLLFKLSQNASLGRHVRPLPYEDKEVEP 145
QY 120 GTLCDVAGCIVNHAAGRPDSLOHVLPLVDORATCNRRTHGDCAITERLMAESNRDSC 179
120 GTLCDVAGCIVNHAAGRPDSLOHVLPLVDORATCNRRTHGDCAITERLMAESNRDSC 179
DB 146 GTLCDVAGCIVNHAAGRPDSLOHVLPLVDORATCNRRTHGDCAITERLMAESNRDSC 205
146 GTLCDVAGCIVNHAAGRPDSLOHVLPLVDORATCNRRTHGDCAITERLMAESNRDSC 205
QY 180 KDSGSGPLVCGGVLGVVTHSGSVCCGKPKGIVTVASVAAIDSVL 226
180 KDSGSGPLVCGGVLGVVTHSGSVCCGKPKGIVTVASVAAIDSVL 226
DB 206 KDSGSGPLVCGGVLGVVTHSGSVCCGKPKGIVTVASVAAIDSVL 252

Query Match 69.7%; Score 844; DB 6; Length 260;
Best Local Similarity 66.2%; Pred. No. 4.5e-82;
Matches 151; Conservative 33; Mismatches 42; Indels 2; Gaps 1;
QY 1 ILGREAHAARPYMASVOLNGAHLCGGVLAQWVLSAARHLEDAADGKQVOLLGAH 58
1 ILGREAHAARPYMASVOLNGAHLCGGVLAQWVLSAARHLEDAADGKQVOLLGAH 58
26 ILGREAHAARPYMASVOLNGAHLCGGVLAQWVLSAARHLEDAADGKQVOLLGAH 85
59 LSQPSPEKLLTYLRAVPHDSQPTIDHLLQLSEKATLGPVAPRLPMQVDRDVA 118
59 LSQPSPEKLLTYLRAVPHDSQPTIDHLLQLSEKATLGPVAPRLPMQVDRDVA 118
DB 86 LSAPPEYKRWTDVQSVVPHGSPDSDLEDDLLFKLSQNASLGRHVRPLPYEDKEVE 145
86 LSAPPEYKRWTDVQSVVPHGSPDSDLEDDLLFKLSQNASLGRHVRPLPYEDKEVE 145
QY 119 GTLCDVAGCIVNHAAGRPDSLOHVLPLVDORATCNRRTHGDCAITERLMAESNRDSC 178
119 GTLCDVAGCIVNHAAGRPDSLOHVLPLVDORATCNRRTHGDCAITERLMAESNRDSC 178
DB 146 GTLCDVAGCIVNHAAGRPDSLOHVLPLVDORATCNRRTHGDCAITERLMAESNRDSC 205
146 GTLCDVAGCIVNHAAGRPDSLOHVLPLVDORATCNRRTHGDCAITERLMAESNRDSC 205
QY 179 KDSGSGPLVCGGVLGVVTHSGSVCCGKPKGIVTVASVAAIDSVL 226
179 KDSGSGPLVCGGVLGVVTHSGSVCCGKPKGIVTVASVAAIDSVL 226
DB 206 KDSGSGPLVCGGVLGVVTHSGSVCCGKPKGIVTVASVAAIDSVL 253

RESULT 8
US-08-944-483-45
: Sequence 45, Application US/08944483
: Patent No. 6232456
: GENERAL INFORMATION:
: APPLICANT: COHEN, MAURICE
: APPLICANT: COLPITTS, TRACEY L.
: APPLICANT: FRIEDMAN, PAULA N.
: APPLICANT: GRAYSON, RANDI N.
: APPLICANT: KLASG, MICHAEL R.
: APPLICANT: RUSSELL, JOHN C.
: APPLICANT: STEWART, KENT D.
: APPLICANT: STROUPE, STEVEN D.
: TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
: TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
: NUMBER OF SEQUENCES: 6
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSSEE: Abbott Laboratories
: STREET: 100 Abbott Park Road
: CITY: Abbott Park
: STATE: IL
: COUNTRY: USA
: CONTACT NO.: 606/4642500
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/944,483
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE: INFORMATION:

ATTORNEY/AGENT INFORMATION:

NAME: Hecker, Cheryl L.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6183.US.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:

LENGTH: 232 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6232456e

US-08-944-483-45

Query Match

ches 100: Similarity 38.9%; Score 471.5; DB 4; Length 232;

Best Local Similarity 41.3%; Pred. No. 2e-42;

Matches 100: Conservative 41; Mismatches 81; Indels 9; Gaps 5;

QY 1 ILGDEAEMAPMPSVOLANGHLCGGVLAPOWVLSAHLCDLDAAGCKVQLGLHSL 60

Db 1 ILGDEAEMAPMPSVOLANGHLCGGVLAPOWVLSAHLCDLDAAGCKVQLGLHSL 60

Db 1 ILGDEAEMAPMPSVOLANGHLCGGVLAPOWVLSAHLCDLDAAGCKVQLGLHSL 60

QY 61 SOPEPSKELDVLRAVPHDSQP-DTIDHLLQLSEKATLGPVRLPQWRDVPAP 119

Db 60 DSPQLT--FHKAATQHPKVPALENDLALLODKYKSTIRPLALSKKROVAA 116

QY 120 GTCDVAGWGVHAGRRDSLOHVLPLVDLWATCNRTTH--DCAITELKCAESNR--R 176

Db 117 GTCSMACWGLTHGGRLSLVELDLQVLTVCNENRWNGLSPSKWCLAADSKDA 176

QY 179 -CKDGGSGPLVC--GVLEGVTSGRVCCNRKKPGIYTRVASYAAWIDSV 226

Db 177 FKCDGGSGPLVCGRVLAGVLSFSRVCITDFRPVATAPVSVIRKV 227

RESULT 9

US-08-944-483-46

Sequence 46, Application US/08944483

Patent No. 6232456

GENERAL INFORMATION:

APPLICANT: COLPITTS, MAURICE

APPLICANT: COLPITTS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GRANADOS, EDWARD N.

APPLICANT: KLAAS, MICHAEL R.

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STEWART, KENT D.

APPLICANT: STEWART, KENT D.

TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

TITLE OF INVENTION: OF THE PROSTATE

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: ILLINOIS

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/944,483

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE: INFORMATION:

ATTORNEY/AGENT INFORMATION:

NAME: Hecker, Cheryl L.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6183.US.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:

LENGTH: 234 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6232456e

US-08-944-483-46

Query Match

ches 100: Similarity 37.7%; Score 456.5; DB 4; Length 234;

Best Local Similarity 41.5%; Pred. No. 7.9e-41;

Matches 95: Conservative 44; Mismatches 81; Indels 9; Gaps 5;

QY 1 ILGDEAEMAPMPSVOLANGHLCGGVLAPOWVLSAHLCDLDAAGCKVQLGLHSL 60

Db 1 ILGDEAEMAPMPSVOLANGHLCGGVLAPOWVLSAHLCDLDAAGCKVQLGLHSL 60

Db 1 ILGDEAEMAPMPSVOLANGHLCGGVLAPOWVLSAHLCDLDAAGCKVQLGLHSL 60

QY 61 SOPEPSKELDVLRAVPHDSQP-DTIDHLLQLSEKATLGPVRLPQWRDVPAP 120

Db 58 TREPTPTQIMLHKEFPYDPAVREGDLQLQTEKAKINKYVTLHLPRGGDDVKPG 117

QY 121 TCGDVAGWGVHAGRRDSLOHVLPLVDLWATCNRTTH--DCAITELKCAESNR--R 176

Db 118 TMCQVAGWGVHAGRRDSLOHVLPLVDLWATCNRTTH--DCAITELKCAESNR--R 176

QY 177 DCKGDSGSGPLVCGLVGGVTSVSG-SRVCGNRKKPGIYTRVA-SYAAMT 223

Db 178 DSCNGDSGSGPLVCGRVLAGVLSFSRVCITDFRPVATAPVSVIRKV 226

RESULT 10

US-07-990-301A-4

Sequence 4, Application US/07990301A

Patent No. 5500346

GENERAL INFORMATION:

APPLICANT: SAVERS, THOMAS; WILTROUT, THERESA A.

APPLICANT: POWERS, JAMES C.; SMYTH, MARK J.; SOMDER, RAYMOND;

APPLICANT: HENDERSON, LOUIS E.

TITLE OF INVENTION: NOVEL SERINE PROTEASE AND

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 345 PARK AVE.

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/990,301A

FILING DATE: 03-DEC-1992

ATTOREY/AGENT INFORMATION:

ATTORNEY/AGENT INFORMATION:

REFERENCE/DOCKET NUMBER: 2025-4055

FILING DATE:

CLASSIFICATION: 212-758-4800

TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO. 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 232
TYPE: AMINO ACID
TOPOLOGY: UNKNOWN
MOLECULE TYPE: PROTEIN
DESCRIPTION: YES
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: HUMAN
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL LINE: LOPEZ LGL LEUKEMIA
ORGANELLE:
US-07-990-301A-4

Query Match 37.48; Score 452.5; DB 1; Length 232;
Best Local Similarity 43.38; Pred. No. 2.1e-40;
Matches 100; Conservative 37; Mismatches 85; Indels 9; Gaps 5;
Qy 1 ILGREGAEAHAPRYMASVOLGNAHLCGGVLAQWYLSAARCLDADQKQVVLGAHSL 60
Db 1 ILGREGAEAHAPRYMASVOLGNAHLCGGVLAQWYLSAARCLDADQKQVVLGAHSL 60
Db 1 ILGREGAEAHAPRYMASVOLGNAHLCGGVLAQWYLSAARCLDADQKQVVLGAHSL 60
Qy 61 SOPEPSKELDVLRAVPHDSQPTDTHDLLLLLQLSKATLGPVRLPMQWVDROPAP 119
Db 60 DSPLT---THKALQIPRYKPVALVQALQDQKQVRLPRLPMQWVDROPAP 119
Qy 120 GTLCDVAGWGIYVNAHGRDPSLOHVLPLVLDRAICNRRTHIDGAIATERLCAESNRD 177
Db 117 GTRCSMGWLGTHGGGSLRVLAEDLDQLVLDTRCNNSRFNGSLSPKVLAAQSLDQA 176
Qy 178 SCGDSGGPLVCG-GVLEGVYTSGRVGNKKGKGIYTRVASYAMTDSV 226
Db 177 PCGDSGGPLVCGGRVLAGVLSFSRVCTDIEKPPVATAVPYVMIRKY 227

RESULT 11
US-07-990-301A-2
Sequence 2; 560134
Patent No. 6263979
GENERAL INFORMATION:
APPLICANT: SATERS, THOMAS; WILTROUT, THERESA A.;
APPLICANT: POWERS, JAMES C.; SMYTH, MARK J.; SOMDER, RAYMOND;
TITLE OF INVENTION: NOVEL SERINE PROTEASE AND
TITLE OF INVENTION: NOVEL SERINE PROTEASE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVE.
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT # 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/990,301A
FILING DATE: 03-DEC-1992
ATTORNEY/AGENT INFORMATION:
FIRM: FELLIS, 728
REGISTRATION NUMBER: 2025-4055
REFERENCE/DOCKET NUMBER: 2025-4055

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 258
TYPE: AMINO ACID
TOPOLOGY: UNKNOWN
MOLECULE TYPE: PROTEIN
DESCRIPTION: NO
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: RAT
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE: CYTOLYTIC GRANULE
CELL LINE: RNK-16 LGL
ORGANELLE:
US-07-990-301A-2

Query Match 36.78; Score 445; DB 1; Length 258;
Best Local Similarity 38.88; Pred. No. 1.5e-36;
Matches 92; Conservative 48; Mismatches 85; Indels 6; Gaps 5;
Qy 1 ILGREGAEAHAPRYMASVOLGNAHLCGGVLAQWYLSAARCLDADQKQVVLGAHSL 60
Db 1 ILGREGAEAHAPRYMASVOLGNAHLCGGVLAQWYLSAARCLDADQKQVVLGAHSL 60
Qy 61 SOPEPSKELDVLRAVPHDSQPTDTHDLLLLLQLSKATLGPVRLPMQWVDROPAP 119
Db 80 HDPOPSGLTETVKAQIKHPVNL-KYENDALALKDGRVPSKNVXPLAPKPKXPAP 138
Qy 120 GTLCDVAGWGIYVNAHGRDPSLOHVLPLVLDRAICNRRTHIDGAIATERLCAESNRD 178
Db 139 GRCSTAGWGIYVNAHGRDPSLOHVLPLVLDRAICNRRTHIDGAIATERLCAESNRD 178
Qy 179 -CKDSGGPLVCG-GVLEGVYTSGRVGNKKGKGIYTRVASYAMTDSV 227
Db 199 PCGDSGGPLVCGGRVLAGVLSFSRVCTDIEKPPVATAVPYVMIRKY 249

RESULT 12
US-09-008-271A-7
Sequence 7
Patent No. 6263979
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Yue, Henry
Guegler, Karl J.
Guegler, Neil C.
Tadokoro, Tomi
Shah, Purvi
TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
REGISTRATION NUMBER: 2025-4055
APPLICATION NUMBER: US/09/008,271A

```

1 FILING DATE: 16-JAN-1998
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: <Unknown>
4 FILING DATE: <Unknown>
5 ATTORNEY/AGENT INFORMATION:
6 NAME: SHEILA
7 REGISTRATION NUMBER: 41,201
8 REFERENCE/DOCKET NUMBER: 8F-0458 US
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: 650-845-0555
11 TELEFAX: 650-845-4166
12 INFORMATION FOR SEQ ID NO: 1:
13 SEQUENCE CHARACTERISTICS:
14 LENGTH: 260 amino acids
15 TYPE: amino acid
16 STRANDEDNESS: single
17 TOPOLOGY: linear
18 IMMEDIATE SOURCE:
19 LIBRARY: COLNOT27
20 COUNTRY: USA
21 GENE: 179846
22 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
23 -008-271A-7
24
25 Query Match
26 Best Local Similarity 36.4%; Score 440.5; DB 4; Length 260;
27 Matches 90; Conservative 44; Mismatches 85; Indels 13; Caps 6;
28
29 QY 1 ILGREAERAHRYPTMASVOLNGAHLGGVLYASQWVLSAAHCLDEADAQGVYVLLGAMSL 60
30 QY 22 LVHGGCKRDTYHPYQAALYTSYSHLLGGVLYHPLVYTAHCKRP----NQQVFLGRLH 77
31 QY 61 SQPEPSKRLYDLVRAVHP---DSQPDYTDHLLQLLSEKATLGPVRLPQWRVDV 120
32 QY 78 RORESSQESSVYPAVHPYDAASHDDQIMLLRLARPAKLSLQIPLP---LEROSCAN 134
33 QY 121 TL-CDVAGAGTGVNHAGRPOSLOHVLVLPVLDORATCNRRTHHGDGAITERL 177
34 QY 135 TFSCHILWG-KTAGDGPDTICANYHLSVRECEHA--YPOITOMNLGAGDEKYGKD 191
35 QY 178 SKCGDSGGLVCGGVLYSEKVTSGSCVRCNKRKPGIYTRVASYAMIDSVL 227
36 QY 192 SCQDGSGLVCGDHLRGLVSMGNIPGSGKPGKGVTVNCRVITWNIQKTI 241
37
38 RESULT 14
39 US-09-070-526-2
40 Sequence 2, Application US/09070526
41 Patent No. 6100059
42 GENERAL INFORMATION:
43 APPLICANT: SOUTHAN, CHRISTOPHER
44 APPLICANT: CLINKENBEARD, HELEN
45 APPLICANT: BURGESS, NICOLA
46 INVENTOR: SOUTHAN, CHRISTOPHER
47 NUMBER OF SEQUENCES: No. 6100059el Compounds
48 TITLE OF INVENTION: PROTEASE AND RELATED DNA COMPOUNDS
49 CORRESPONDENCE ADDRESS:
50 ADDRESSEE: RATNER & PRESTIA
51 STREET: P. O. BOX 980
52 CITY: VALLEY FORGE
53 STATE: PA
54 COUNTRY: USA
55 ZIP: 19224
56 COMPUTER READABLE FORM:
57 MEDIUM TYPE: Diskette
58 OPERATING SYSTEM: DOS
59 SOFTWARE: FASTSEQ for Windows Version 2.0
60 CURRENT APPLICATION DATA:
61 APPLICATION NUMBER: US/09/070,526
62 FILING DATE: 30-APR-1998
63 CLASSIFICATION:
64 PROTEIN INFORMATION:
65 APPLICATION NUMBER: GB 9711952.3
66 FILING DATE: 9-JUN-1997
67 APPLICATION NUMBER: EP 97309646.4
68 FILING DATE: 1-DEC-1997
69 ATTORNEY/AGENT INFORMATION:
70 NAME: PRESTIA, PAUL F
71 REGISTRATION NUMBER: 23,031
72 INVENTOR: SOUTHAN, CHRISTOPHER
73 TELECOMMUNICATION INFORMATION:
74 TELEPHONE: 610-407-0700
75 TELEFAX: 610-407-0701
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
81
```

TELEX: 846169
; INFORMATION: SEQ ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 260 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-070-326-2

Search completed: May 1, 2003, 22:09:27
Job time : 30 secs

Query Match 35.88; Score 433.5; DB 3; Length 260;
Best Local Similarity 39.58; Pred. No. 2.6e-38;
Matches 90; Conservative 42; Mismatches 83; Indels 13; Gaps 6;
QY 1 ILGREA-EAHARYMASVOLNAGHLCGGVLYARQWYLSAAHCLDEADAGKQVVLGAHSL 60
DB 33 VLGGHECPHSPQNALFGQGLLGGVLYGGWVLTAAHCKRP-----KYTVLGDHSL 88
QY 61 SQPEPSKRLDYLVRAVPHP---DSQPTDIDHLLQLSEKATLGPVRLPQWRVDV 117
DB 89 QNKDGPQDEIPVQSGIPHCYNSDSVDHNDHMLLQLRDQASIGSKVKPI--SLADHCT 146
QY 118 ARTGLDVAGKIWNHAGRR-PDSIQHVLVPLDRATCNRTHDGAITERLMCAESNR- 175
DB 147 QPCQKTVSGCTVTSRENFPUTLNCARVIFPKQKC--EDAYPQGITGCAVCASSRG 204
QY 176 RDSCKGSGPLVCGGVLEGVTSRGVCGNRKKKPGITRVASYAAMI 223
DB 205 ADTCQDGSGLVCDGALQGITSMGSDPCGRSKDPGVYTNICRYLDWI 252

RESULT 15
US-09-386-642-13
; Sequence 13, Application us/09386642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Barron, Andrew
; INVENTOR: Barron, Andrew
; APPLICANT: QLT Jensen
; INVENTOR: QLT Jensen
; TITLE OF INVENTION: Zymogen Activation System
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/386.642
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 13
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusin gene
; OTHER INFORMATION: with homo sapien serine protease catalytic domain
US-09-386-642-13

Query Match 34.88; Score 421; DB 4; Length 288;
Best Local Similarity 38.68; Pred. No. 6.4e-37;
Matches 90; Conservative 43; Mismatches 86; Indels 14; Gaps 7;
QY 1 ILGREA-EAHARYMASVOLNAGHLCGGVLYARQWYLSAAHCLDEADAGKQVVLGAHSL 59
DB 52 IYGTGCEPHSPQNALFGQGLLGGVLYGGWVLTAAHCKRP---KYTVLGDHSL 107
QY 60 LSQPEPSKRLDYLVRAVPHP---DSQPTDIDHLLQLSEKATLGPVRLPQWRVDV 116
DB 108 LQNKDGPQDEIPVQSGIPHCYNSDSVDHNDHMLLQLRDQASIGSKVKPI--SLADHCT 165
QY 117 VARTGLDVAGKIWNHAGRR-PDSIQHVLVPLDRATCNRTHDGAITERLMCAESNR 175
DB 166 TQPCQKTVSGCTVTSRENFPUTLNCARVIFPKQKC--EDAYPQGITGCAVCASSRG 223
QY 176 -RDSCKGSGPLVCGGVLEGVTSRGVCGNRKKKPGITRVASYAAMIDSVL 227
DB 224 GADTCQDGSGLVCDGALQGITSMGSDPCGRSKDPGVYTNICRYLDWIKKII 276

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sv model

Run on: May 1, 2003, 22:01:05 : Search time 76 Seconds
(without alignments)
399,732 Million cell updates/sec

Title: US-09-821-255-2

Perfect score: 1211

Sequence: 1 ILGRCAGHAPRYMASVOL.....KPGIYIRVSYAAMDSVLA 228

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

i number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1202	99.3	250	11 AAR05772	Human adipisin gene
2	1201	99.2	250	11 AAR05421	Human adipisin/D en
3	1098.5	90.7	212	20 AAY28590	Human Factor D, H
4	1056	87.2	233	20 AAY28591	Pig Factor D, S
5	456.5	37.7	262	8 AAE71672	Human serine prote
6	435.1	37.7	264	18 AAW27496	Human H83-22 secre
7	435.1	37.7	264	18 AAW27496	Amino acid sequenc
8	447.5	37.0	232	15 AAR55758	Human serine protease
9	447	36.9	283	20 AAY28640	Human Factor D, H
10	447	36.9	283	21 AAY9371	Human PRO1549 (UNQ

11	447	36.9	283	21 AAY70036	Human Protease and
12	447	36.9	283	22 AAU29159	Human PRO polypept
13	447	36.9	283	22 AAU02497	Human secreted pro
14	447	36.9	283	22 AAB66120	Protein of the inv
15	445	36.7	258	15 AAE57577	Human Granzym L pr
16	445	36.7	258	15 AAE57577	Human Granzym L pr
17	440.5	36.4	260	20 AAY06438	Human Protease HUP
18	437.5	36.1	260	20 AAY14744	Human PRO322 prote
19	437.5	36.1	260	20 AAY32852	Human serine prote
20	437.5	36.1	260	20 AAY03220	Amino acid sequenc
21	437.5	36.1	260	21 AAB21322	Human neuropilin.
22	437.5	36.1	260	21 AAB21322	Human neuropilin.
23	437.5	36.1	260	21 AAE51131	Human PRO322 (UNQ2
24	437.5	36.1	260	21 AAE51131	Human PRO322 (UNQ2
25	437.5	36.1	260	22 AAG23373	Novel human diaph
26	437.5	36.1	260	22 AAU12067	Human PRO322 polyp
27	437.5	36.1	260	22 AAB53087	Human angiogenesis
28	437.5	36.1	260	23 ABB94548	Human angiogenesis
29	437.5	36.1	260	23 ABB84852	Human PRO322 prote
30	437.5	36.1	275	21 AAY21311	Human PRO322, Hom
31	437.5	36.1	305	20 AAY32853	Human secreted prote
32	437.5	36.1	305	21 AAB37985	Human secreted prote
33	437.5	36.1	315	23 ABR41332	Human ovarian anti
34	433.5	35.8	244	15 AAR44532	Zyme APP-cleaving
35	433.5	35.8	244	18 AAR22985	Protease M, a nove
36	433.5	35.8	244	19 AAW51006	Human zyme, Homo
37	433.5	35.8	244	19 AAW51006	Human zyme, Homo
38	433.5	35.8	260	20 AAB77703	Engraulis japonicu
39	432.5	35.7	240	23 ABB04644	Human zyme, Homo
40	431.5	35.6	239	21 AAB21310	Human trypsin doma
41	424	35.0	249	23 AAE21441	Mouse protease-rel
42	421.5	34.8	253	20 AAY08025	Fusion gene with h
43	421	34.8	288	21 AAB36482	Amino acid sequenc
44	419.5	34.6	286	21 AAB36482	Human secreted prote
45	419.5	34.6	280	17 AAW10594	Human recombinant

ALIGNMENTS

RESULT 1
AAR05772
ID AAR05772 standard; protein; 250 AA.
AC AAR05772:
XX 05-NOV-1990 (first entry)
DE Human adipisin gene product from the clone phg31.
DT
DE Human adipisin; complement D; obesity; lupus erythematosis;
KW
KW human; complement D; obesity; lupus erythematosis;
KW
KW rheumatoid arthritis; ds.
XX
XX MO9006365-A.
XX
XX 14-JUN-1990.
XX
XX 21-NOV-1989; 89MC-0005374.
XX
XX 30-NOV-1988; 88US-0277963.
XX
XX (BETH-) BETH ISRAEL HOSPITA.
XX (DANA-) DANA-FABER CANCER INST.
XX (METH-) METABOLIC BIOSYSTEMS INC.
XX
XX Filler JS, Spiegelman BM, Rosen HM, White RT;
XX WPI: 1990-209777/27.
XX N-PSDB; AAQ05149.
XX
XX Human protein having adipisin and complement D activity -
XX used in diagnosis and control of metabolically caused obesity
XX and for treating infection.
XX

XX Disclosure: : p: English.

XX this raised to the gene product may be used in diagnosis of
XX metabolically obese, obesity. The product may also be used to
XX treat and prevent obesity and bacterial, viral, parasitic and
XX neoplastic cell infection.

XX Sequence 250 AA:

Query Match 99.3%; Score 1202; DB 11; Length 250;
Best Local Similarity 99.1%; Pred. No. 1.4e-105;
Matches 226; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILGGRERAEHARPYMASVOLNGAHLGCGVLVABQWVLSAAHCLDEADAAGKVOVLLGAHSL 60

Db 23 ILGGRERAEHARPYMASVOLNGAHLGCGVLVABQWVLSAAHCLDEADAAGKVOVLLGAHSL 82

QY 61 SQEPESKRLYDYLRAVPHFDSQPTIDHLLQLLQSEKATLGPVRLPQWRVDVAFG 120

Db 83 SQEPESKRLYDYLRAVPHFDSQPTIDHLLQLLQSEKATLGPVRLPQWRVDVAFG 142

QY 121 TLCDVAGGVVNHAGRRPDSQLVPLVDRATCNRRTHDCAITERLMCAESNRDSC 180

Db 143 TLCDVAGGVVNHAGRRPDSQLVPLVDRATCNRRTHDCAITERLMCAESNRDSC 202

QY 181 GDSGGPLVCGGVLEGVVTSGSRVCCNKKPKGITTRVASTAAWIDSVLA 228

Db 203 GDSGGPLVCGGVLEGVVTSGSRVCCNKKPKGITTRVASTAAWIDSVLA 250

RESULT 2

AA05421

ID AAR05421 standard; protein; 250 AA.

AC AAR05421;

DT 30-JUL-1990 (first entry)

XX Human adipisin/D encoded by a cDNA.

XX CAT, hybrid protein; Human adipisin/D.

XX Homo sapiens.

XX W09001540-A.

XX 22-FEB-1990.

XX 19-AUG-1989; 89WO-US03417.

XX 11-AUG-1988; 88US-0231224.

XX (CALB-) CALIF BIOTEC INC.

XX Hilliker S.; White R;

XX WPI; 1990-083499/11.

XX N-PSDB: AAQ03566.

XX Heterologous protein expression on prokaryotic host -

XX the 3' end of the cDNA was inserted into a pMT vector to

XX stably express hybrid protein.

XX Example: Fig 10: 67pp; English.

XX When inserted into an expression vector, pTRPCAT 72,

XX the construct gave 10-15% levels of fusion protein upon

XX induction in W3110 cells.

XX See also AAQ03557 to AAQ05366; and AAQ04767.

XX Sequence 250 AA:

Query Match 99.3%; Score 1201; DB 11; Length 250;
Best Local Similarity 99.1%; Pred. No. 1.4e-105;
Matches 226; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILGGRERAEHARPYMASVOLNGAHLGCGVLVABQWVLSAAHCLDEADAAGKVOVLLGAHSL 60

Db 23 ILGGRERAEHARPYMASVOLNGAHLGCGVLVABQWVLSAAHCLDEADAAGKVOVLLGAHSL 82

QY 61 SQEPESKRLYDYLRAVPHFDSQPTIDHLLQLLQSEKATLGPVRLPQWRVDVAFG 120

Db 83 SQEPESKRLYDYLRAVPHFDSQPTIDHLLQLLQSEKATLGPVRLPQWRVDVAFG 142

QY 121 TLCDVAGGVVNHAGRRPDSQLVPLVDRATCNRRTHDCAITERLMCAESNRDSC 180

Db 143 TLCDVAGGVVNHAGRRPDSQLVPLVDRATCNRRTHDCAITERLMCAESNRDSC 202

QY 181 GDSGGPLVCGGVLEGVVTSGSRVCCNKKPKGITTRVASTAAWIDSVLA 228

Db 203 GDSGGPLVCGGVLEGVVTSGSRVCCNKKPKGITTRVASTAAWIDSVLA 250

RESULT 3

AA028590

ID AAY28590 standard; Protein; 212 AA.

AC AAY28590;

XX 26-NOV-1999 (first entry)

XX Human Factor D.

XX Factor D; alternative complement pathway; serine protease; asthma;

XX C3/C5 convertase; C3/C5 convertase; tissue damage; Alzheimer's disease;

XX multiple sclerosis; acute myocardial infarction; Crohn's disease;

XX transplant rejection; systemic lupus erythematosus; inflammation.

XX Homo sapiens.

XX W09942133-A1.

XX 26-AUG-1999.

XX 19-FEB-1999; 99WO-US03566.

XX 20-FEB-1998; 98US-0075328.

XX (TANO-) TANOX INC.

XX Fung MSC, Sun BNC, Sun CRY;

XX WPI; 1999-527424/44.

XX N-PSDB: AA206673.

XX Factor D inhibitors useful for treating pathological inflammation

XX and autoimmune diseases

XX Example 1: Page 93-94; 99pp; English.

XX This sequence is human factor D. Factor D is a highly specific serine

XX protease essential for activation of the alternative complement pathway.

XX The plasma concentration of factor D in human is very low making it a

XX poor candidate for antibody production. The human factor D gene has

XX activation of the alternative complement pathway. Factor D is used to

XX generate inhibitors of complement activation which bind factor D at a

XX molar ratio of at least 1.5:1 and less than 80:1 (inhibitor:factor D).

XX Factor D was used to design a monoclonal antibody Mab 166-32 which binds

XX to factor D and blocks its ability to activate complement. The

XX anti-factor D molecules can function to inhibit in vivo complement

XX activation and/or the alternative complement pathway and inflammatory

XX responses which include the activation of the alternative complement

XX macrophages, neutrophils, platelets and mast cells, oedema, and tissue

XX damage. The inhibitors of complement activation may be used to treat

XX diseases or conditions that are mediated by excessive or uncontrolled


```

XX FH /label= mat_protein
XX EP245051-A.
XX 11-NOV-1987.
XX PF 01-MAY-1987; 87EP-0303945.
XX PR 31-DEC-1986; 86US-0948248.
XX PR 06-MAY-1986; 86US-0846248.
XX PR 08-MAY-1986; 86US-0861221.
XX (STED ) LELAND STANFORD JR UNIV.
XX PA Weissman IL, Gershenfeld HK;
XX PI WPI: 1987-315213/45.
XX DR N-PSDB; AAM71407.
XX
XX New pure human serine protease and fragments - used as labels in
XX assays and for prodn. of antibodies for passive immunisation
XX against immune disorders.
XX
XX Disclosure; Fig 1; 7pp; English.
XX
XX Amino acid homology within the active enzyme to the mouse,protease is
XX 71% with 77% at the DNA level. The overall homology is 72% when the
XX complete coding region and the 3' UTR are included. The amino acids
XX of the charge-relay system are His41, Asp86 and Ser184. The acidic
XX residue Asp178 determines substrate specificity for Lys or Arg.
XX The Hsp is produced by activated killer cells. The enzyme acts in
XX conjunction with other components of a killer cell to provide
XX cytolytic capability.
XX
XX Sequence 262 AA:
XX
XX Query Match 37.7%; Score 456.5; DB 8; Length 262;
XX Best Local Similarity 41.5%; Pred. NO. 4.7e-35;
XX Matches 95; Conservative 44; Mismatches 81; Indels 9; Gaps
XX
XX QY 1 ILGCREAARHPYMASVOLGNGAHLCGGVIVASQWYLSAAHCLDEADGKQVVLGASL 60
XX DB 29 IIGNEVTPHSPRYWLLSDRTICAGALLKMDWLTFAHC---NANKRSQVILGASL 85
XX QY 61 SPPPSKRLVDYLRVPHDSQPTDIDHLLQLSEKATLGPVRLPQWRDVRAP 120
XX DB 86 FREETPKIIMLKVEFPYPCVDPATRGSDLKLLQTEKAKINKVTYLIHLPRKGDVKG 145
XX Q 121 TLCDVAGVIVHAGRRPDSIQHVLVLPVLDRTCNRRTHD--DGAITERLKAESNR--R 176
XX DB 146 THCVAGVGRTHNAGSSQSLREVTITIDRKNVDNRHNFNPVIGRWVYAGSLRGR 205
XX QY 177 DSCKDGSGPLGVGVLEQVVTSG-SRVCGRNKKPGYITRYA-SYAAKRT 223
XX DB 206 DSCNGLSGSPLLCEVPRGVTSEGLNCKGDPGRPGVYILLSKHLNMI 254
XX
XX RESULT 6
XX AAM27296 standard; Protein; 264 AA.
XX
XX AC AAM27296;
XX
XX DT 19-MAY-1998 (first entry)
XX
XX DE Human H83-22 secreted protein.
XX
XX Secreted protein; anti-inflammatory; immune stimulant; suppressant;
XX human cytokine; autoimmune disease; regulator; activin, inhibin;
XX H83-22.
XX
XX OS Homo sapiens.

```

```

XX FH Location/Qualifiers
XX FT 1..284
XX FT 1..284 H83-22
XX FT /note= "secreted protein"
XX
XX W09740151-A2.
XX
XX 30-OCT-1997.
XX
XX 14-APR-1997; 97MO-US06125.
XX
XX 10-JAN-1997; 97US-0781226.
XX
XX 19-APR-1996; 96US-0635311.
XX
XX (GRW ) GENETICS INST INC.
XX
XX Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;
XX Spaulding V;
XX
XX WPI: 1997-535835/49.
XX
XX N-PSDB; AAT91308.
XX
XX Nucleic acids encoding novel secreted proteins - useful e.g. as an
XX anti-inflammatory, immune stimulant or suppressor, etc.
XX
XX Claim 24; Page 65-66; 81pp; English.
XX
XX This sequence represents a novel secreted protein, H83-22, which is
XX isolated from a clone, H83-22 (ATCC 98029), derived from a human PBMC
XX cDNA library. The nucleic acid and encoded H83-22 secreted protein can be
XX used as probes for the detection of H83-22 mRNA or for the detection of
XX markers for cells primers (e.g. for nutrition (as C, N or
XX carbohydrate source), as a cytokine for cell proliferation and
XX differentiation activity, as immune stimulants or suppressors e.g.
XX for viral, bacterial or fungal infections, for autoimmune diseases such
XX as multiple sclerosis or systemic lupus erythematosus, to regulate
XX hematopoiesis, for tissue growth, as an activin or inhibin, or having
XX immunomodulatory activity, as an anti-inflammatory, receptor/ligand,
XX anti-inflammatory or tumour inhibitory activities.
XX
XX Sequence 264 AA:
XX
XX Query Match 37.7%; Score 456.5; DB 18; Length 264;
XX Best Local Similarity 42.7%; Pred. NO. 4.7e-35;
XX Matches 100; Conservative 40; Mismatches 85; Indels 9; Gaps
XX
XX QY 1 ILGCREAARHPYMASVOLGNGAHLCGGVIVASQWYLSAAHCLDEADGK-VQVILGAS 59
XX DB 27 IIGCEVSPHSPRPMASIQYGGHHVCGGVLDIPQWLTAAHCQYRPTKQGSPTVILGAS 86
XX QY 60 LSQPEPSKRLYDVLVAPHSQPTDIDHLLQLSEKATLGPVRLPQWRDVRAP 119
XX DB 87 LSRNEASQKTELEKFFIPFSVTSQSDNDMLVRLQTAANKVHKLW-HISKSTLSRS 145
XX QY 120 QTLCDVAGVIVHAGRRP-DSIQHVLVLPVLDRTCNRRTHDG--ALTERLKA--ESN 174
XX DB 146 GTKCKVTGNGADPDSLRPSDTLREVTYVLSRKLCSQSYNGDPFTKQVCAGDAKG 205
XX
XX QY 175 RDSCKDGSGLPVGVGVLEQVVTSGSRVCGNRKKPGYITRYA-SYAAIDSVL 227
XX DB 206 QNDSCKDGSGLPVGVGVLEQVVTSGSRVCGNRKKPGYITRYA-SYAAIDSVL 258
XX
XX RESULT 7
XX AAM27254
XX
XX ID AAM27254 standard; Protein; 284 AA.
XX
XX AC AAM27254;
XX
XX DT 23-APR-2002 (first entry)
XX
XX DE Amino acid sequence of novel human protease #53.

```

XX human; protease; cancer; immune-related disorder; cardiovascular disease;
 KW diseases; metabolic disorder; inflammatory disorder;
 KW nervous system; psychomotoric disorder; inflammatory disorder;
 KW hypertension; psychotic disorder; neurological disorder; dyskinesia;
 KW viral infection; human immunodeficiency virus; HIV; non-viral infection;
 KW ocular disease; cytostatic; enzyme.

XX Homo sapiens.
 XX W0200200860-A2.
 XX 03-JAN-2002.
 XX 26-JUN-2001; 2001WO-US20171.
 XX 26-JUN-2000; 2000US-214047P.
 XX (SUGB) SUGEN INC.

XX Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
 XX Charydzak G;
 XX NPI: 2002-138913/18.
 XX N-PSDB; ARK31795.
 XX Nucleic acids encoding novel human proteases, useful for useful for
 PT treating diseases and disorders such as cancers, immune-related
 PT diseases and disorders, cardiovascular disease (e.g. restenosis) and
 PT inflammatory disorders -
 XX Claim 6; Fig 2R; 313pp; English.

XX The present invention relates to the isolation of novel human
 CC proteases, and the nucleic acids encoding them. The sequences of
 CC the invention are useful for treating diseases and disorders such as
 CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders
 CC (e.g. inflammatory diseases and asthma), cardiovascular diseases
 CC diseases, metabolic disorders (e.g. diabetes, obesity), central or
 CC disorders (e.g. rheumatoid arthritis and psoriasis), sexual dysfunction,
 CC peripheral nervous system diseases, migraines, pain, sexual dysfunction,
 CC mood disorders, attention disorders, cognition disorders, hypotension,
 CC hypertension, psychotic disorders, neurological disorders
 CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.
 CC The invention also relates to methods for treating viral
 CC infections caused by human immunodeficiency virus (HIV), non-viral
 CC infections such as ocular disease (e.g. glaucoma) and macular
 CC degeneration. AAU82702-AAU82760 represent the novel human proteases of
 CC the invention.

XX Sequence 284 AA:
 Query Match 37.28; Score 451; DB 23; Length 284;
 Best Local Similarity 42.74; Pred. No. 1.7e-34;
 Matches 100; Conservative 34; Mismatches 88; Indels 12; Gaps 6;
 QY 1 IIGGREAARVYMASVQLAGHLCGGVLAQWVLSAHL--EDAADGKQVVLGAH 58
 Db 35 IIGGVTTHSRPPYMASVRFQGHGCGFLHARWVSNAHCFSHRLDRTG--LVVLGAH 92
 QY 59 SLSQPEKRLDYLRVAVPHDSQDPDLIDHLLQLLSEKATLGPV--RPLPMQVRDAP 116
 Db 93 VLSTAEPQTQVGDITLTHPDYHPWTHANDICLLQNLGNSVLPAGVLLGPGR-AP 151
 QY 117 VAPGTLCQVAGVYVNHAGRPDSLOHVLVPLVDRTACNRRTHHGATLTERLAC---AES 173
 Db 152 PTATGTCRKGAGVGFDEELPGLMEAKVYVDPVCN--SSMKGHLLTLMCTRGDS 209
 QY 174 NRDCSKGDSGGPLVCGVGLGVYVSSGRVGNKKPGVITRVASYANTDSVL 227
 Db 210 HRRGFCSDSGPLVCHNRAHGLVFSFGLWCQIDKPTPDVYTVQSAFVAKIWDV 263

RESULT 8
 AAR55758
 XX AAR55758 standard; Protein; 232 AA.
 XX AAR55758;
 AC AAR55758;
 DT 17-NOV-1994 (first entry)
 XX Serine protease Met-ase.
 DE Serine protease; RNK Met-1, cytolytic granule; leukemia;
 KW large granular lymphocyte; RNK-16; Met-ase; tumor; diagnosis;
 KW DNA probe; hybridization; Lopez.
 XX Homo sapiens.
 XX W09412647-A.
 XX 09-JUN-1994.
 PD 03-DEC-1993; 93WO-US11736.
 PF 03-DEC-1992; 92US-0990301.
 PR (US) US DEPT HEALTH & HUMAN SERVICES.
 XX Henderson LE, Powers JC, Sayer T, Smyth MJ, Sowder B;
 PI Wiltrout TA;
 XX WPI: 1994-200275/24.
 DR N-PSDB; AAQ66910.
 XX New serine protease with Met-ase activity - isolated from
 PT lymphocyte granules, also related DNA vectors, transformed cells
 PT and antibodies.
 XX Disclosure: Page 38-39; 57pp; English.
 XX A new 30 kDa serine protease, RNK Met-1, was isolated from the
 CC cytolitic granules of cells RNK-16 large granular leukemia
 CC cells. cDNA was obtained from a rat RNK-16 lambda-otII library and
 CC sequenced (AAQ66909). The mature protein has the sequence given in
 CC AAR55757. The corresponding sequences for the human serine protease,
 CC Met-ase, obtained from Lopez large granular lymphocyte leukemia
 CC cells are given in AAQ66910 and AAR55758. DNA encoding RNK Met-1
 CC can be used as a probe to detect the enzyme in tumor biopsy samples.
 XX Sequence 232 AA;

Query Match 37.04; Score 447.5; DB 15; Length 232;
 Best Local Similarity 42.44; Pred. No. 2.9e-34;
 Matches 98; Conservative 38; Mismatches 86; Indels 9; Gaps 5;
 QY 1 IIGGSEAEARVYMASVQLAGHLCGGVLAQWVLSAHL--EDAADGKQVVLGAH 60
 Db 1 IIGGVTTHSRPPYMASVRFQGHGCGFLHARWVSNAHCFSHRLDRTG--LVVLGAH 92
 QY 61 SLSQPEKRLDYLRVAVPHDSQDP--DTIDHLLQLLSEKATLGPVRLPMPQVRDAP 119
 Db 60 DSPGLT--FHIAAIOHPKPKVPAVLFQALLQDQGVKFSPTIRLPLSKHQVAA 116
 QY 120 GTICQVAGVYVNHAGRPDSLOHVLVPLVDRTACNRRTHHGATLTERLACSENKRS- 178
 Db 117 GTRCSMAQGLVHGGGSLRVLAELDQLVDITNCKNSRFNGSVSPNVLCAADSDQA 176
 QY 179 -CKDGGSGPLVC--GGVLGVYVSSGRVGNKKPGVITRVASYANTDSV 226
 Db 177 PCKDGGSGPLVCGKRGVLAGVLSFSRSRVSTDTFKPPVATAVAPVYSWKV 227
 RESULT 9
 AAY28640

	98US-0101474.	PR	17-NOV-1998:	98US-0106925.	XX
	98US-0101475.	PR	18-NOV-1998:	98US-0106926.	XX
	98US-0101476.	PR	18-NOV-1998:	98US-0106927.	XX
	98US-0101477.	PR	18-NOV-1998:	98US-0106928.	XX
	98US-0101478.	PR	18-NOV-1998:	98US-0106929.	XX
	98US-0101479.	PR	18-NOV-1998:	98US-0106930.	XX
	98US-0101480.	PR	18-NOV-1998:	98US-0106931.	XX
	98US-0101481.	PR	18-NOV-1998:	98US-0106932.	XX
	98US-0101482.	PR	18-NOV-1998:	98US-0106933.	XX
	98US-0101483.	PR	18-NOV-1998:	98US-0106934.	XX
	98US-0101484.	PR	18-NOV-1998:	98US-0106935.	XX
	98US-0101485.	PR	18-NOV-1998:	98US-0106936.	XX
	98US-0101486.	PR	18-NOV-1998:	98US-0106937.	XX
	98US-0101487.	PR	18-NOV-1998:	98US-0106938.	XX
	98US-0101488.	PR	18-NOV-1998:	98US-0106939.	XX
	98US-0101489.	PR	18-NOV-1998:	98US-0106940.	XX
	98US-0101490.	PR	18-NOV-1998:	98US-0106941.	XX
	98US-0101491.	PR	18-NOV-1998:	98US-0106942.	XX
	98US-0101492.	PR	18-NOV-1998:	98US-0106943.	XX
	98US-0101493.	PR	18-NOV-1998:	98US-0106944.	XX
	98US-0101494.	PR	18-NOV-1998:	98US-0106945.	XX
	98US-0101495.	PR	18-NOV-1998:	98US-0106946.	XX
	98US-0101496.	PR	18-NOV-1998:	98US-0106947.	XX
	98US-0101497.	PR	18-NOV-1998:	98US-0106948.	XX
	98US-0101498.	PR	18-NOV-1998:	98US-0106949.	XX
	98US-0101499.	PR	18-NOV-1998:	98US-0106950.	XX
	98US-0101500.	PR	18-NOV-1998:	98US-0106951.	XX
	98US-0101501.	PR	18-NOV-1998:	98US-0106952.	XX
	98US-0101502.	PR	18-NOV-1998:	98US-0106953.	XX
	98US-0101503.	PR	18-NOV-1998:	98US-0106954.	XX
	98US-0101504.	PR	18-NOV-1998:	98US-0106955.	XX
	98US-0101505.	PR	18-NOV-1998:	98US-0106956.	XX
	98US-0101506.	PR	18-NOV-1998:	98US-0106957.	XX
	98US-0101507.	PR	18-NOV-1998:	98US-0106958.	XX
	98US-0101508.	PR	18-NOV-1998:	98US-0106959.	XX
	98US-0101509.	PR	18-NOV-1998:	98US-0106960.	XX
	98US-0101510.	PR	18-NOV-1998:	98US-0106961.	XX
	98US-0101511.	PR	18-NOV-1998:	98US-0106962.	XX
	98US-0101512.	PR	18-NOV-1998:	98US-0106963.	XX
	98US-0101513.	PR	18-NOV-1998:	98US-0106964.	XX
	98US-0101514.	PR	18-NOV-1998:	98US-0106965.	XX
	98US-0101515.	PR	18-NOV-1998:	98US-0106966.	XX
	98US-0101516.	PR	18-NOV-1998:	98US-0106967.	XX
	98US-0101517.	PR	18-NOV-1998:	98US-0106968.	XX
	98US-0101518.	PR	18-NOV-1998:	98US-0106969.	XX
	98US-0101519.	PR	18-NOV-1998:	98US-0106970.	XX
	98US-0101520.	PR	18-NOV-1998:	98US-0106971.	XX
	98US-0101521.	PR	18-NOV-1998:	98US-0106972.	XX
	98US-0101522.	PR	18-NOV-1998:	98US-0106973.	XX
	98US-0101523.	PR	18-NOV-1998:	98US-0106974.	XX
	98US-0101524.	PR	18-NOV-1998:	98US-0106975.	XX
	98US-0101525.	PR	18-NOV-1998:	98US-0106976.	XX
	98US-0101526.	PR	18-NOV-1998:	98US-0106977.	XX
	98US-0101527.	PR	18-NOV-1998:	98US-0106978.	XX
	98US-0101528.	PR	18-NOV-1998:	98US-0106979.	XX
	98US-0101529.	PR	18-NOV-1998:	98US-0106980.	XX
	98US-0101530.	PR	18-NOV-1998:	98US-0106981.	XX
	98US-0101531.	PR	18-NOV-1998:	98US-0106982.	XX
	98US-0101532.	PR	18-NOV-1998:	98US-0106983.	XX
	98US-0101533.	PR	18-NOV-1998:	98US-0106984.	XX
	98				

PR 01-DEC-2000; 2000MO-US2578.
 PI 20-DEC-2000; 2000MO-US34956.
 XX (GETH) GENENTECH INC.
 PA Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 XX WPI; 2001-602746/68.
 DR N-PSDB; AAS46060.
 XX
 XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 PT presence of tumours, such as prostate and breast tumours, in mammals and
 PT screen for modulators of the compounds -
 XX Claim 11; Fig 272; 774pp; English.
 XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
 CC The PRO polypeptides and their associated nucleic acids can be used to
 CC detect the presence of a tumour in a mammal by comparing the level of
 CC expression of a PRO polypeptide in a test sample of cells from the animal
 CC and a control sample of normal cells, whereby a higher level of
 CC expression in the test sample indicates the presence of a tumour in the
 CC animal. The animal can be a cow, horse, sheep, pigs, goats
 CC and rabbits but are preferably human. The polypeptides can be used to
 CC stimulate tumour necrosis factor (TNF) alpha release from human blood
 CC when contacted with it. A specific polypeptide can be used to stimulate
 CC the proliferation or differentiation of chondrocyte cells. The PRO
 CC proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC animals. The nucleic acids can be used as probes for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.
 XX Sequence 283 AA;
 SO
 Query Match 36.9%; Score 447; DB 22; Length 283;
 Best local similarity 42.3%; Pred. No. 4e-34; Indels 12; Gaps 6;
 Matches 99; Conservative 39; Mismatches 88;
 OY 1 ILGREAERAPYMASVOLNGLCGVILVARQWLSAAHCL-EDAADGKVOVLGAH 58
 DB 34 ILGGHEVTPHPRYMASYREGGHHGGGFLLRARWYVSAHCTSHRLDRTG-LVVLGAH 91
 OY 59 SLQSPFSKRLVDLRAVPHDSQPTIDHLLLOLSEKATLGPAY-RPUPQWRD 116
 DB 92 VLSTARTQNGIDALTPDTHPDHNTANDCLLNKSSALGPAVLGLRRA-RP 150
 OY 117 VARGTLDVAGNGVHVAHGRDSQIQLVFLVDLRATCNRRHGDATLRLMC--AES 173
 DB 151 PTAGTRCRVAGNGVSDPEELPGLEAKRVILDPDVCN--SSWKGHTILMLCTRSDS 208
 OY 174 NRDSCKGSGRACVCGVLEGVTSQGVQNRKKGLITRWASVAMIDSVL 227
 DB 209 HRGFCSDSGPLVLCRNBAHGLVSGLCGCDPKTDVYTSVFANVTDVW 262
 RESULT 13
 AAU02497
 D AAU02497 standard; Protein; 283 AA.
 AC AAU02497;
 XX
 DT 07-SEP-2001 (first entry)
 XX Human secreted protein TANGO 298.
 DE Human secreted protein; TANGO 298.
 KW Human secreted protein; TANGO 298; clone JYHmal18f02; chromosome 19p13;
 KW complement factor D; alternative complement pathway;
 KW deficiency; hereditary; autosomal recessive; human; human adipsin;
 KW obesity; diabetes; blood and hematopoietic associated disorder;
 KW cardiovascular disorder; inflammatory disorder; immune disorder.
 XX

OS Homo sapiens.
 XX Key
 XX Peptide
 FT /label= Signal_peptide
 FT /note= "N-myristoylation site"
 FT Modified-site
 FT 28..33
 FT Protein
 FT 32..283
 FT Domain
 FT 34..238
 FT 34..238 Trypsin_domain
 FT Modified-site
 FT 50..52
 FT /note= "Protein kinase C phosphorylation site"
 FT Modified-site
 FT 55..60
 FT /note= "N-myristoylation site"
 FT Active-site
 FT 70..75
 FT /note= "Histidine active site common to serine
 FT protease, trypsin family"
 FT Modified-site
 FT 77..79
 FT /note= "Protein kinase C phosphorylation site"
 FT Modified-site
 FT 77..80
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site
 FT 94..97
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site
 FT 110..113
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site
 FT 129..132
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site
 FT 144..147
 FT /note= "Asn is N-glycosylated"
 FT Modified-site
 FT 166..169
 FT /note= "Amidation site"
 FT Modified-site
 FT 169..179
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site
 FT 189..192
 FT /note= "N-myristoylation site"
 FT Modified-site
 FT 191..193
 FT /note= "Asn is N-glycosylated"
 FT Modified-site
 FT 208..210
 FT /note= "Protein kinase C phosphorylation site"
 FT Modified-site
 FT 230..241
 FT /note= "Protein kinase C phosphorylation site"
 FT Modified-site
 FT 263..266
 FT /note= "N-myristoylation site"
 FT Modified-site
 FT 275..277
 FT /note= "cAMP and cGMP dependent protein kinase
 FT phosphorylation site"
 FT Modified-site
 FT 277..279
 FT /note= "Protein kinase C phosphorylation site"
 PN WO200130831-A1.
 XX
 PD 03-MAY-2001.
 XX
 XX 27-OCT-2000; 2000MO-US29797.
 XX 27-OCT-1999; 98US-0417796.
 PR 17-MAY-2000; 2000US-0572275.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 PA
 XX Fraser CC, Hodge MR;
 PI WPI; 2001-300479/31.
 DR N-PSDB; AAS04269.
 XX
 XX New nucleic acid molecule encoding type II transmembrane proteins
 PT useful for treating immune related disorders -
 PT Claim 23; Fig 10; 137pp; English.
 XX The present sequence representing a novel human secreted protein
 CC TANGO 298, which is a member of the complement factor D gene
 CC cDNA library. The gene for TANGO 298 maps to chromosome 19p13.
 CC TANGO 298 shows sequence homology to human adipsin (complement

PT helper sequence -

PS Disclosure; Fig 6; 67pp; German.

XX
XX
CC This invention describes a novel method for producing (M1) one or more
CC serine proteases and/or one or more domains of one or more serine
CC proteases in a prokaryotic host. (M1) comprises: (1) a serine protease
CC and/or one or more domains of them are expressed comprising an N-terminal
CC Met-Y; (2) renaturing the recombinant peptide with a dipeptide sequence
CC N-terminal sequencing. Y = any amino acid, especially proline. The
CC invention is useful for recombinantly producing serine proteases for the
CC potential use in the treatment of inflammatory and autoimmune disorders.
CC The recombinantly produced active proteins comprises only the
CC catalytically active domain, giving them a specific function with natural
CC or artificial specificity.

XX Sequence 283 AA;

Very Match 36.8%; Score 446; DB 22; Length 283;
Best Local Similarity 42.3%; Pred. No. 5e-34;
Matches 99; Conservative 35; Mismatches 88; Indels 12; Gaps 6;
QY 1 ILGRENAAHPYMASVOLNCAHLGCVLYAEQVYLSAHLG--EDAAQCKVULLGAA 58
Db 34 IIGCHEVTHSRPYMASVRFQGHGCGFLLKAAWVYSAARCFSHRLNTG--LVVLGAH 91
QY 59 SLSPQPSKRLYDLRAVPDPDPTIDHLLQLSEKATLQPAV--RLEPWQRVDRD 116
Db 92 VLSTAEPTQVFGIDALTTHPTPHMTANDICLLKNGSAVLGPVGLLPGRA--RP 150
QY 117 VAPGTLCDVAGCIVNHAGRPDSLOHVLPLVDLRATCNERTHDGATPELMC---AES 173
Db 151 PTVGTRCRVAGMGFYSDFEELPGLMEAKRVYLDPCVN--SSMKGHLTLTLCITRSGDS 208
QY 174 NHRDCKDGGGLVCGVLEGVVTSSGRVCGNKKPGIYTRVASYAAMDSVL 227
Db 209 HRRGFTSADGGPLVCNENANGVTSFSLWCDPRTDPTQVQKRFAMINDVV 262

Search completed: May 1, 2003, 22:07:13
Job time : 79 secs

A:Accession: A26105
 A:Molecule type: mRNA
 A:Residues: 1-259 <IN>
 A:Cross-references: GB:X49883; PID:CAA28378.1; PID:9581866
 C:Comment: Human complement factor D is synthesized primarily in cells of the macrophage lineage. However, expression of the murine homology is specific to adipose tissue.
 A:Gene: 28K
 A:Introns: 19/1; 71/2; 120/3; 206/3
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: adipose tissue; signal alternative splicing; hydrolase; serine proteinase
 F:1-20/Domin: signal sequence status predicted <SIG>
 F:24-250/Domin: propeptide status predicted <PRO>
 F:26-249/Domin: signal sequence status predicted <SIG>
 F:26-249/Domin: trypsin homology <TRY>
 F:51-67,149-215,180-196,205-230/disulfide bonds: status predicted
 F:56,115,209/Active site: His, Asp, Ser status predicted
 Query Match 59.7% Score 844.5; DB 1; Length 259;
 Best Local Similarity 66.5%; Pred. No. 2.7e-69;
 Matches 151; Conservative 33; Mismatches 42; Indels 1; Gaps 1;

QY 1 ILGGEAAHARYMASVOLNAGCGVLAQWVLSAAHCLDEDA--DSKVQLLGANS 59
 DB 26 ILGGEAAHARYMASVOLNAGCGVLAQWVLSAAHCLDEDA--DSKVQLLGANS 85
 QY 60 LSQPEPKRLDYLVRAVPHDSQPTIDHDLQLLQSEKATIGAPVRLPMDVROVAP 119
 DB 86 LSAPPEVRYDYQVSYVPHDSQPTIDHDLQLLQSEKATIGAPVRLPMDVROVAP 145
 QY 120 GTLGDVAGNGVNIAGRPDSIQHVLVPLDRAVCNRRTHDCAITELRCAESNRDSC 179
 DB 146 GTLGDVAGNGVYTHAGRPDVLHQLVRSINRRTHDCAITELRCAESNRDSC 205
 QY 180 KDGSGSLVCGGVLEVVTSQSVGRNKKPKGPIITRVASYAAMDSV 226
 DB 206 RDGSGSLVCGGVLEVVTSQSVGRNKKPKGPIITRVASYAAMDSV 252

RESULT 5
 A55634
 A:Accession: A55634
 A:Molecule type: precursor - human
 A:Residues: 24-Feb-1995 <Accession>
 C:Date: 24-Feb-1995 <Accession>
 C:Accession: A55634
 R:Blat, D.; Fluk, T.; Obermaier-Skrobanek, B.; Zimmer, M.; Wekerle, H.; Lichter, P.; Jørgensen, C.G.
 Genomics 24, 445-450, 1994
 A:Title: The human Net-ase gene (GZMM): structure, sequence, and close physical linkage
 A:Reference number: A55634; MUID:95229142; PMID:7713495
 A:Accession: A55634
 A:Status: precursor
 A:Molecule type: DNA
 A:Residues: 1-249 <PL>
 A:Cross-references: GB:L36922
 A:Gene: GDB:GZMM
 A:Cross-references: GDB:387366; OMIM:600311
 C:Keywords: trypsin; trypsin homology
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; serine proteinase
 F:26-249/Domin: trypsin homology <TRY>
 Query Match 38.8% Score 469.5; DB 2; Length 249;
 Best Local Similarity 43.4%; Pred. No. 3.3e-35;
 Matches 99; Conservative 41; Mismatches 79; Indels 9; Gaps 5;

QY 1 ILGGEAAHARYMASVOLNAGCGVLAQWVLSAAHCLDEDA--DSKVQLLGANS 60
 DB 26 ILGGEAAHARYMASVOLNAGCGVLAQWVLSAAHCLDEDA--DSKVQLLGANS 85
 QY 61 LSQPEPKRLDYLVRAVPHDSQPTIDHDLQLLQSEKATIGAPVRLPMDVROVAP 119
 DB 85 DSGELT---PHIKAIQDPRKVPVLAENDLLQLQSKVTSKTRTALPASKROVAA 141

QY 120 GTLGDVAGNGVNIAGRPDSIQHVLVPLDRAVCNRRTHDCAITELRCAESNRDSC 178
 DB 142 GTGCSKAGNLTHQGGRLSVLRDELQVLTGHCNSRFRNLSLSSVAVCLAKDQA 201
 QY 179 -CKGDSGGPLVC--CGVLGVWTSQSVGRNKKPKGPIITRVASYAAMT 223
 DB 202 PKGDSGGPLVC--CGVLGVWTSQSVGRNKKPKGPIITRVASYAAMT 249

RESULT 6
 A31372
 A:Accession: A31372
 A:Molecule type: precursor [validated] - human
 A:Residues: 07-Jul-1995 <Accession>
 C:Date: 07-Jul-1995 <Accession>
 C:Accession: A31372; PID:930684
 R:Gershenfeld, H.K.; Hershenberger, R.J.; Shows, T.B.; Weissman, I.L.
 Proc. Natl. Acad. Sci. U.S.A. 85, 1184-1188, 1988
 A:Title: Cloning and chromosomal assignment of a human cDNA encoding a T cell-
 A:Reference number: A31372; MUID:88125000; PMID:3257574
 A:Accession: A31372
 A:Status: precursor
 A:Molecule type: protein
 A:Residues: 1-262 <GSR>
 A:Cross-references: GB:M8737; NID:gl84022; PID:AAAS2647.1; PID:930684
 R:Hamed, A.; Lowrey, D.M.; Lichtenheld, M.; Podack, E.R.
 J. Immunol. 141, 3142-3147, 1988
 A:Title: Characterization of three serine esterases isolated from human IL-2 activat
 A:Reference number: A92823; MUID:8909886; PMID:3262682
 A:Accession: A30525
 A:Status: precursor
 A:Molecule type: protein
 A:Residues: 29-40 <CHAM>
 R:Kraehenbuehl, O.; Rey, C.; Jenne, D.; Lanzavecchia, A.; Groscurth, P.; Carrel, S.; J. Biol. Chem. 263, 13215-13222, 1988
 A:Title: Characterization of granzymes A and B isolated from granules of cloned huma
 A:Reference number: A92824; MUID:89035468; PMID:3263427
 A:Accession: A30526
 A:Status: precursor
 A:Molecule type: protein
 A:Residues: 29-36
 R:Poe, M.; Bennett, C.D.; Biddison, W.E.; Blake, J.T.; Norton, G.P.; Rodkey, J.A.; S
 J. Biol. Chem. 263, 13215-13222, 1988
 A:Title: Human cytotoxic lymphocyte proteinase. Its purification from granules and the
 A:Reference number: A28943; MUID:88330824; PMID:3047119
 A:Accession: A28943
 A:Status: precursor
 A:Molecule type: protein
 A:Residues: 1-46; R:48-49, 'X', 51, 'X', 53 <PO>
 A:Experimental source: CTL line Q31
 A:Note: Details from reference A39027 (see entry A61021) suggest this protein is gra
 C:Genetics:
 A:Gene: GDB:GZMA; CTLA3; HFSP
 A:Cross-references: GDB:120601; OMIM:140050
 A:Map position: 5q11-5q12
 C:Keywords: glycoprotein; hydrolase; trypsin homology
 F:1-26/Domin: signal sequence status predicted <SIG>
 F:27-28/Domin: propeptide status predicted <PRO>
 F:29-36/Product: granzyme A status experimental <MPT>
 F:29-36/Domin: trypsin homology <TRY>
 F:54-70/4th-21st/Active site: His, Asp, Ser (covalent) status predicted
 F:170/Binding site: carbohydrate (Asn) (covalent) status predicted
 Query Match 37.7% Score 456.5; DB 1; Length 262;
 Best Local Similarity 41.5%; Pred. No. 5.4e-34;
 Matches 95; Conservative 44; Mismatches 81; Indels 9; Gaps 5;

QY 1 ILGGEAAHARYMASVOLNAGCGVLAQWVLSAAHCLDEDA--DSKVQLLGANS 60
 DB 29 ILGGEVTHRYLVRAVPHDSQPTIDHDLQLLQSEKATIGAPVRLPMDVROVAP 85
 QY 61 LSQPEPKRLDYLVRAVPHDSQPTIDHDLQLLQSEKATIGAPVRLPMDVROVAP 120
 DB 86 TREPTQKLVVLRKPEVPCYPDARTREDGLQLQTEKAKINRYVTTLHLKPKGDVQK 145

QY 121 TLCDVAGNGVNHAGRRPDSQVHLLPVLDRATCNRRTH--DGAITERLMCAENR--R 176
 Db 146 TPCQVANGHTNSAKSHSLREVLIIIDRKVNDNRHYNFVIGNWYCAQSGRGR 205
 QY 177 DSKNGSGGGLVGVGVGVGVTSV--SRVCGNKKKPGIYTRVA--SYAAMI 223
 Db 206 DSCNCGSGPLCCGVRGFRVTSGLKNCQDGRPGVQVILLSKHLNMI 254
 RESULT 7
 S65663
 granzyme 3 (EC 3.4.21.-) precursor - human
 N:Alternate names: cytotoxic T-lymphocyte proteinase 3
 C:Species: Homo sapiens (man)
 C:Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 22-Jun-1999
 P:Accession: U05686; GI:86417; SCD45
 P:Release: N.M. Ocasio, S. Schmitz, B.F. Fiers Lett. 364, 268-271, 1995
 A:Title: Cloning of cDNA for human granzyme 3.
 A:Reference number: S65617; NID:95278340; PMID:7758581
 A:Accession: S65663
 A:Molecule type: mRNA
 A:Cross-references: EMBL:U26174; NID:9829637; PIDN:AA475478.1; PID:gb29638
 A:Accession: S65617
 A:Molecule type: protein
 A:Residues: 27-42 <R>
 R:Nameed, A.; Lowrey, D.M.; Lichtenheld, M.; Podack, E.R.
 J. Immunol. 141, 3142-3147, 1988
 A:Reference number: A93823; NID:8509466; PMID:3262682
 A:Accession: C30525
 A:Molecule type: protein
 A:Residues: 27-42 <HAM>
 C:Genetics:
 A:Gene: GDB:GMK; PRSS
 A:Cross-references: GDB:605612; OMIM:600704
 A:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; serine proteinase
 F:1-36/Domain: signal sequence #status predicted <SIG>
 F:27-264/Product: granzyme 3 #status experimental <MAT>
 F:27-254/Domain: trypsin homology <TRY>
 Query Match 37.7%; Score 456.5; DB 2; Length 264;
 Best Local Similarity 42.7%; Pred. No. 5.4e-34;
 Matches 100; Conservative 40; Mismatches 85; Indels 9; Gaps 7;
 C 1 ILIGREAEAHAPRYMSVOLGANGHLCGGVLADEQVYLSAARCLDEADAGK--VOVLLGAHS 59
 Db 21 IIGREAEAHAPRYMSVOLGANGHLCGGVLADEQVYLSAARCLDEADAGK--VOVLLGAHS 60
 QY 60 LQSPDSKRLDYLVRAVPHDPDSQPTDIHLLQLSEKATGLPAVRLPMQVRDVPAP 119
 Db 87 LSKNEASKQTLKEKTPSPVTSQPTDINLVKLTAAKLKHKML--HRSKTSLSRS 145
 QY 120 GTLCDVAGNGVNHAGRRP--DSQVHLLPVLDRATCNRRTHG--AITEKLMA--ESN 174
 Db 146 GTCKVTCGAGDPSRLVSDTLREVVYVLSKLCNSQSYNGDIFTKMVCAGAGK 205
 QY 175 RRSCKGSGGGLVGVGVGVGVTSV--SRVCGNKKKPGIYTRVA--SYAAMISVL 227
 Db 206 QKDSCKGSGGGLVGVGVGVGVTSV--SRVCGNKKKPGIYTRVA--SYAAMISVL 254
 RESULT 8
 I56220
 trypsin 2 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 06-Dec-1996 #sequence_revision 26-Jul-1996 #text_change 22-Jun-1999
 P:Accession: I56220
 R:Sayers, T.J.; Wilttrout, T.A.; Smyth, M.J.; Ottaway, K.S.; Pillaro, A.M.; Sowder, R.; Hd

J. Immunol. 152, 2289-2297, 1994
 A:Title: Purification and cloning of a novel serine protease, RNK-Tryp-2, from the gr
 A:Reference number: I56220; NID:9419809; PMID:8135042
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-258 <RES>
 A:Cross-references: GB:L19694; NID:9487878; PIDN:AAA42057.1; PID:9487879
 C:Superfamily: trypsin; trypsin homology
 F:25-248/Domain: trypsin homology <TRY>
 Query Match 37.2%; Score 451; DB 2; Length 258;
 Best Local Similarity 41.0%; Pred. No. 1.7e-33;
 Matches 98; Conservative 42; Mismatches 77; Indels 22; Gaps 9;
 QY 1 ILIGREAEAHAPRYMSVOLGANGHLCGGVLADEQVYLSAARCLDEADAGK--VOVLLGAHS 60
 Db 26 IIGREAEAHAPRYMSVOLGANGHLCGGVLADEQVYLSAARCLDEADAGK--VOVLLGAHS 83
 QY 61 SQPDSKRLDYLVRAVPHDPDSQPTDIHLLQLSEKATGLPAVRLPMQVRDVPAP 120
 Db 84 SKNPFKQTLKEKTPSPVTSQPTDINLVKLTAAKLKHKVQ--LHURSKYLRQ 140
 QY 121 PLQVAGNGVNHAGRRP-----DSQVHLLPVLDRATCNRRTH--HGAITERLMCAE 172
 Db 141 TNCQVTCGAGDPSRLVSDTLREVVYVLSKLCNSQSYNGDIFTKMVCAG 195
 QY 173 SNR--RDSCKGSGGGLVGVGVGVTSV--SRVCGNKKKPGIYTRVA--SYAAMISVL 228
 Db 196 DRGKDSCKGSGGGLVGVGVGVTSV--SRVCGNKKKPGIYTRVA--SYAAMISVL 253
 RESULT 9
 A45161
 serine proteinase (EC 3.4.21.-) RNK-Met-1 - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 P:Accession: A45161; GI:10417; T. Traub, J.A.; Ottaway, K.S.; Sowder, R.; Henderson, L.E.
 J. Biol. Chem. 267, 24418-24425, 1992
 A:Title: Purification and cloning of a novel serine protease, RNK-Met-1, from the gra
 A:Reference number: A45161; NID:93077530; PMID:1447189
 A:Accession: A45161
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Cross-references: GB:L05175; NID:9206691; PIDN:AAA42056.1; PID:9206692
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; serine proteinase
 F:21-245/Domain: trypsin homology <TRY>
 Query Match 36.7%; Score 445; DB 2; Length 258;
 Best Local Similarity 39.8%; Pred. No. 5.9e-32;
 Matches 92; Conservative 48; Mismatches 85; Indels 6; Gaps 5;
 QY 1 ILIGREAEAHAPRYMSVOLGANGHLCGGVLADEQVYLSAARCLDEADAGK--VOVLLGAHS 60
 Db 21 IIGREAEAHAPRYMSVOLGANGHLCGGVLADEQVYLSAARCLDEADAGK--VOVLLGAHS 79
 QY 61 SQPDSKRLDYLVRAVPHDPDSQPTDIHLLQLSEKATGLPAVRLPMQVRDVPAP 119
 Db 80 HDPDQGLTFLFKNGHNGNL--KYENDLALQDGRVPSKYNKPLAPRPPAPAE 138
 QY 120 GTLCDVAGNGVNHAGRRPDSQVHLLPVLDRATCNRRTHGCA--TERLMCAENRDS-- 178
 Db 139 GSRCKTAGWGLTHQGLAKSLQELDLRLDTRMCRNRSRWYVLTDSMLCAAGKGA 198
 QY 179 --CKDSCKGSGGGLVGVGVGVTSV--SRVCGNKKKPGIYTRVA--SYAAMISVL 227
 Db 199 PKGDSCKGSGGGLVGVGVGVTSV--SRVCGNKKKPGIYTRVA--SYAAMISVL 249
 RESULT 10

569370 duodenase - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 24-Jul-1998 sequence_revision 24-Jul-1998 #text_change 21-Nov-1998
C:Accession: S66570
C:Length: 227
C:Keywords: hydrolase; serine proteinase
A:Title: Duodenase, a new serine protease of unusual specificity from bovine duodenal mucosa
A:Reference number: 569370; MUID:95172076; PMID:7867649
A:Accession: 569370
A:Molecule type: protein
A:Residues: 1-226 <ZAM>
C:Superfamily: trypsin; trypsin homology
F:1-210/Domain: trypsin homology <TRY>
Query Match 35.94; Score 434.5; DB 2: Length 226;
Best Local Similarity 40.94; Pred. No. 4.5e-32;
Matches 95; Conservative 44; Mismatches 78; Indels 15; Gaps 6;
1 ILGREGAARHAPYASV--QLNG-AHLGCGVGVAVQWLSAAHCLDEADAGKVOYLGA 57
2 ILGCHENARHAPYASV--LFTSGSHICGFLVDRDVPVTAHCL-----GSHVTLGA 55
3 ILGCHENARHAPYASV--LFTSGSHICGFLVDRDVPVTAHCL-----GSHVTLGA 55
4 ILGCHENARHAPYASV--LFTSGSHICGFLVDRDVPVTAHCL-----GSHVTLGA 55
5 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
6 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
7 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
8 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
9 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
10 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
11 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
12 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
13 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
14 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
15 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
16 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
17 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
18 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
19 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
20 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
21 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
22 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
23 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
24 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
25 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
26 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
27 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
28 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
29 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
30 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
31 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
32 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
33 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
34 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
35 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
36 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
37 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
38 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
39 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
40 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
41 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
42 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
43 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
44 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
45 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
46 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
47 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
48 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
49 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
50 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
51 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
52 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
53 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
54 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
55 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
56 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
57 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
58 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
59 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
60 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
61 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
62 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
63 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
64 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
65 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
66 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
67 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
68 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
69 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
70 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
71 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
72 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
73 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
74 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
75 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
76 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
77 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
78 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
79 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
80 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
81 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
82 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
83 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
84 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
85 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
86 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
87 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
88 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
89 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
90 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
91 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
92 HSLGSPESKRLDYLRVAPHPDSQPTIDHLLQLLSEKATGLGPAVPLPWQRVDROV 117
93 HSL

```

Qy 61 SOPEPSKELVYLRAVHP--DSDPTDIDHLLQLLSEKATLGPVRLPQWRVDV 117
Db 89 QSRDPEQETQVANS--DHCPCWNSFNEDSHIDNLSANLQKQKVPVLANCPKY 148
Qy 118 APTGLCDVAGMGVGNHAGR-PDSLQHLVPLVDLRATCNRRTHDGAITERLMCA 175
Db 149 --GQKICISQWGTVPQENPNTLCAEVIYSKNCERA--YPGKITECHWCYCSNG 204
Qy 176 RSCGDSGGPLVCGVLEGVTSGRVGNKRRKIITRVASTAAVDSVL 227
Db 180 SCQDGGSGPLVCGVLEGVTSGRVGNKRRKIITRVASTAAVDSVL 256

RESULT 14
TRYPTR
A:Title: (EC 3.4.21.4) precursor - pig (tentative sequence)
M:Accession: A90368
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 31-Mar-2000
C:ession: A90641; A90368; A00947
R:ries, M.; Rovey, M.; Guidoni, A.; Desnuelle, P.
Biochim. Biophys. Acta 69, 115-129, 1963
A:Title: Su ile trypsinogen and la trypsin de porc.
A:Ref: C. Desnuelle, 1963
A:Accession: A90641
A:Molecule type: protein
A:Residues: 1-10 <CHA>
P:Hermoson, M.A.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 12, 3146-3153, 1973
A:Title: Determination of the amino acid sequence of porcine trypsin by sequenator analy
A:Ref: M. A. Hermoson, L. H. Ericsson, K. A. Walsh, 1973
A:Accession: A90368; NUID:7325892; PMID:4736935
A:Molecule type: protein
A:Residues: 9-231 <HER>
A:Note: at position 20, Ile and Val occur alternatively
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase; zym
F:1-4/Domain: activation site; trypsin homology <21A>
F:9-231/Domain: trypsin homology <21A>
F:9-231/Product: trypsin #status experimental <MAT>
F:9-224/Domain: trypsin homology <TRY>
F:15-145,33-49,117-218,124-191,156-170,181-205/Disulfide bonds: #status predicted
F:48-92,185/Active site: His, Asp, Ser #status predicted
F:60,62,65,70/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Watch 34.4% Score 416.5; DB 1; Length 231;
Best Local Similarity 37.2%; Pred. No. 2e-30;
..ches 86; Conservative 50; Mismatches 82; Indels 13; Gaps 7;

```

```

Qy 1 ILGREAENAHPTMASVOLNGAHLCGGVLAQWVLSAAHCLDAADCKVYVLLGAHSL 60
Db 9 IYGVITCAINSITFVSLN--SGSHFGCGSLGNSQWYSMAICYK----SRIQVRAGEHNI 63
Qy 61 SOPEPSKELVYLRAVHPDSDPTDIDHLLQLLSEKATLGPVRLPQWRVDV 120
Db 64 DVLGNEQFVINAAKITIHENFNGTDLNDMLIKLSSPATLSNRVATVSLPR--SCAAG 121
Qy 121 TLCDVAGMGVGNHAGR-PDSLQHLVPLVDLRATCNRRTHDGAITERLMCAE--SNRRD 177
Db 122 DELLSGWGTWSSGSSITPLQACVPLVLSGSC--KSSFPQTQNMICVGVLEQCD 179
Qy 178 SCQDGGSGPLVCGVLEGVTSGRVGNKRRKIITRVASTAAVDSVL 228
Db 180 SCQDGGSGPLVCGVLEGVTSGRVGNKRRKIITRVASTAAVDSVL 256

```

```

RESULT 15
P:5061
A:Title: (EC 3.4.21.78) precursor, splice form 2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-May-2000
C:Accession: B45061; A47590

```

```

R:Hershtberger, R.J.; Gershenfeld, H.K.; Weissman, I.L.; Su, L.
J. Biol. Chem. 267, 25488-25493, 1992
A:Title: Genomic organization of the mouse granzyme A gene. Two mRNAs encode the same
A:Reference number: A45061; NUID:9304270; PMID:1460043
A:Accession: B45061; A47590
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <HER>
A:Cross-references: GB:I01429; GB:L00631; NID:91321600; PIDN:AAA9897.1; PID:91321602
R:Hershtberger, R.J.; Gershenfeld, H.K.; Weissman, I.L.
J. Biol. Chem. 267, 25488-25493, 1992
A:Title: Genomic organization of the mouse granzyme A gene. Two mRNAs encode the same
A:Reference number: A45061; NUID:9304270; PMID:1460043
A:Accession: B45061; A47590
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 12-257 <GER>
A:Cross-references: GB:M13226
C:Species: Mus musculus (house mouse)
C:Keywords: hydrolase; serine proteinase
F:26-249/Domain: trypsin homology <TRY>

```

```

Query Watch 34.4% Score 416.5; DB 2; Length 257;
Best Local Similarity 37.5%; Pred. No. 2.3e-30;
Matches 87; Conservative 49; Mismatches 87; Indels 9; Gaps 5;

```

```

Qy 1 ILGREAENAHPTMASVOLNGAHLCGGVLAQWVLSAAHCLDAADCKVYVLLGAHSL 60
Db 26 IIGCDTVVPHSRPYMALIKLSSMTICAGALTEKNVLTAAHC---NVGKRSKFIILGAHSI 82
Qy 61 SOPEPSKELVYLRAVHPDSDPTDIDHLLQLLSEKATLGPVRLPQWRVDV 120
Db 83 NK--EPQIITVTKATPFVPCVDEYTRSDQVPLVLAANKATVNRVAILPLPKGDDVKEG 141
Qy 121 TLCDVAGMGVGNHAGR-PDSLQHLVPLVDLRATCNRRTHD--GATTERLMCAESNR--R 176
Db 142 TRCVAGMGVGNHAGR-PDSLQHLVPLVDLRATCNRRTHD--GATTERLMCAESNR--R 176
Qy 177 SCQDGGSGPLVCGVLEGVTSGRVGNKRRKIITRVASTAAVDSVL 227
Db 202 DSCNGDSGSPLLCDGILHGITSFGGCKDHRNFGYVITFSLQKHLNKKIM 253

```

Search completed: May 1, 2003, 22:08:51
Job time : 20 secs

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 22:03:35 ; Search time 25 Seconds

Title: (partial alignments)
Perfect score: 1211 378.264 Willion G811 updates/sec

Sequence: 1 ILGGRARHARHTMASVOL.....KPGIYTRVASTAAWIDSVLA 228

Scoring table: BLOSUM62

Gapped 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues 112892

#1 number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt-40.*

Prad. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1211	100.0	253	1	CPAD_HUMAN
2	1056	87.2	259	1	P00746 homo sapien
3	869.5	71.8	263	1	P51779 sus scrofa
4	844.5	69.7	259	1	P30338 rattus norv
5	827.5	67.8	263	1	P03953 mus musculu
6	457.5	37.8	263	1	P31124 homo sapien
7	457.5	37.8	263	1	P31124 homo sapien
8	456.5	37.7	262	1	P12544 homo sapien
9	456.5	37.7	264	1	P49863 homo sapien
10	451	37.2	258	1	P49863 rattus norv
11	437.5	36.7	258	1	O03238 rattus norv
12	434.5	35.9	260	1	P80219 bos taurus
13	434.5	35.9	260	1	O03238 homo sapien
14	431.5	35.6	260	1	O03760 ovis arie
15	427	35.3	231	1	MCT3_SHEEP
16	426.5	35.2	238	1	P35033 salmo salar
17	425	35.1	245	1	P80931 ovis arie
18	419.5	34.6	260	1	O61955 mus musculu
19	418	34.5	277	1	O00761 homo sapien
20	416.5	34.4	277	1	P00761 sus scrofa
21	416.5	34.4	277	1	P00761 mus musculu
22	416	34.4	247	1	P13878 homo sapien
23	415	34.3	261	1	P20151 homo sapien
24	413.5	34.1	247	1	P06872 canis fami
25	413.5	34.1	261	1	P07288 homo sapien
26	413	34.1	231	1	P35032 salmo salar
27	411.5	34.0	261	1	P28293 mus musculu
28	409.5	33.8	245	1	P19799 xenopus lae
29	409.5	33.8	245	1	P19799 xenopus lae
30	409.5	33.8	256	1	O94215 homo sapien
31	406	33.5	246	1	P32821 rattus norv
32	406	33.5	246	1	P32822 rattus norv
33	406	33.5	249	1	P21842 canis fami

RESULT 1	CPAD_HUMAN	STANDARD	PRT	253 AA.
ID	CPAD_HUMAN			
AC	P00746			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DT	03-JUN-2002 (Rel. 41, Last annotation update)			
DE	Complement factor D (EC 3.4.21.46) (C3 convertase activator)			
DE	(Properdin factor D) (Adipsin).			
GN	DF			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606			
SM	NCBI_TaxID=9606			
SE	SEQUENCE OF 8-253 FROM N.A.			
RX	MEDLINE=92250520; PubMed=1374388;			
RA	White R.T., Damm D., Hancock N., Rosen B.S., Lowell B.B., Usher P.,			
RA	Flier J.S., Spiegelman B.M.;			
RT	"Human adipsin is identical to complement factor D and is expressed			
RT	at high levels in adipose tissue.";			
RT	(2) Biol. Chem. 267:9210-9213(1992).			
RP	SEQUENCE FROM N.A.			
RA	Flier J.S., Spiegelman B.M., Rosen B.M.;			
RL	Patent number WO9006365, 14-JUN-1990.			
RN	[3]			
RX	SEQUENCE OF 26-252.			
RA	MEDLINE=5000441; PubMed=6383466;			
RX	Volanakis J.E., M.M.;			
RT	"Amino acid sequence of human D of the alternative complement			
RT	pathway.";			
RL	Biochemistry 23:2482-2486(1984).			
RN	[4]			
RX	PARTIAL SEQUENCE OF 26-252.			
RX	MEDLINE=84108950; PubMed=6361133;			
RX	Thompson R.A., Gagnon R., Reid K.M.;			
RT	"Partial amino acid sequence of human factor D of the complement system.			
RT	Similarity in sequence between factor D and proteases of non-plasma			
RT	origin.";			
RL	FEBS Lett. 166:347-351(1984).			
RN	[5]			
RX	PARTIAL SEQUENCE OF 26-61 AND 194-220.			
RX	MEDLINE=84256515; PubMed=6821372; B.M.;			
RX	Volanakis J.E., Gagnon R., Reid K.M.;			
RT	"Factor D of the alternative pathway of human complement.			
RT	Purification, alignment and N-terminal amino acid sequences of the			
RT	major cyanogen bromide fragments, and localization of the serine			
RT	residue at the active site.";			
RL	Biochem. J. 187:863-874(1980).			
RN	[6]			
RX	PARTIAL SEQUENCE OF 26-82.			
RX	MEDLINE=84256515; PubMed=6821372; B.M.;			
RX	Volanakis J.E., Gagnon R., Reid K.M.;			
RT	"Partial amino acid sequence of human factor D.homology with serine			
RT	proteases.";			

ALIGNMENTS

Proc. Natl. Acad. Sci. U.S.A. 77:11116-11119(1980).

[7] PARTIAL SEQUENCE OF 26-78.

RR MEDLINE-8105486; Pubmed-6776531.

RR Davis A.E. III:

RR Active site amino acid sequence of human factor D.*;

RR [80]. Natl. Acad. Sci. U.S.A. 77:4398-43942(1980).

RR [80].

RR X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).

RR MEDLINE-91118317; Pubmed-8289289.

RR Narayana S.V., Narayana M., El-Kabbani O., Kilpatrick J.M., Moore D.,

RR Chen X., Bugb C.E., Volanakis J.E., Delucas L.J.:

RR "Structure of human factor D. A complement system protein at 2.0-A

RR resolution." J. Mol. Biol. 235:695-708(1994).

RR [9].

RR X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).

RR MEDLINE-96025834; Pubmed-7592653;

RR Kim S., Narayana S.V., Volanakis J.E.:

RR "Crystal structure of a complement factor D mutant expressing

RR enhanced catalytic activity." J. Mol. Biol. 240:295-305(1995).

RR F.

RR -I- FUNCTION: FACTOR D CLEAVES FACTOR B WHEN THE LATTER IS COMPLEXED

CC WITH FACTOR C3B, ACTIVATING THE C3BB COMPLEX, WHICH THEN BECOMES

CC THE C3 CONVERTASE OF THE ALTERNATE PATHWAY. ITS FUNCTION IS

CC HOMOLOGOUS TO THAT OF C1S IN THE CLASSICAL PATHWAY.

CC -I- CATALYTIC ACTIVITY: Cleaves component factor B (Arg-1-Lys) when in

CC complex with C3b or with cobra venom factor (CVF).

CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -I- FUNCTIONAL SIMILARITY: SIMILAR TO C1S IN THE FEATURE TABLE.

CC REF. 3 SEQUENCE HAD A PEPTIDE IN THE WRONG ORDER AND ANOTHER ONE

CC MISSING.

CC

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation on

CC the European bioinformatics Institute. There are no restrictions on its

CC usage by non-profit institutions as long as its content is in no way

CC distorted and the source is acknowledged. For more information on this

CC entry requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@sib-sib.ch)

CC

DR EMBL; M4526; AAA35527.1; AUT_INIT.

DR PIR; M40197; DBHU.

DR FDB; 1DPP; 25-FEB-96.

DR CASP; 1DPP; 1-AUG-96.

DR PDB; 1DSU; 17-AUG-96.

DR MEROPS; S01.191.-.

DR Gene; HGNC:2771; DF.

DR MIM; 134350.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00089; trypsin.1.

DR PRINTS; SM00722; CHYMOTRYPSIN.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

CC Complement alternate pathway; Plasma; Hydrolase; Serine protease;

CC Zymogen; Glycoprotein; Signal; 3D-structure.

CC SIGNAL: 1 20

CC PROPEP 21 253

CC ACT_SITE 26 253

CC ACT_SITE 66 66

CC ACT_SITE 112 112

CC ACT_SITE 208 208

CC DISULFID 51 67

CC DISULFID 148 214

CC DISULFID 179 195

CC DISULFID 204 229

CC DISULFID 235 253

CC CONFLICT 35 35

CC CONFLICT 40 40

CC CONFLICT 49 49

CC

CC ACTIVATION PEPTIDE (POTENTIAL).

CC ACTIVATION PEPTIDE D.

CC CHARGE RELAY SYSTEM.

CC CHARGE RELAY SYSTEM.

CC CHARGE RELAY SYSTEM.

CC

CC I -> F (IN REF. 1).

CC H -> F (IN REF. 6).

CC H -> V (IN REF. 6).

CC H -> E (IN REF. 4 AND 5).

```

52 G -> A (IN REF. 1 AND 2).
59 G -> R (IN REF. 1 AND 2).
63 S -> T (IN REF. 4).
73 S -> G (IN REF. 4).
86 HSLS -> TSLP (IN REF. 3).
88 HS -> ST (IN REF. 4).
94 MISSING (IN REF. 4).
96 Q -> G (IN REF. 4).
136 Q -> G (IN REF. 4).
178 TGNRTHDHGATE -> KRLYDVL (IN REF. 4).
243 S -> T (IN REF. 3).
243 S -> H (IN REF. 4).
250 S -> H (IN REF. 4).
250 MISSING (IN REF. 4).
250 MISSING (IN REF. 4).
253 AA; 27004 MW; B053587/B05356AD CRC64;
SEQUENCE 253 AA; 27004 MW; B053587/B05356AD CRC64;
Score Match 100.0%; Score 1211; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.9e-107; Indels 0; Gaps 0;
Matches 228; Conservative 0; Mismatches 0;
QY 1 ILGGEAEAHAPRYMASVOLGNHLCGGVAVABOMVLASNCLEDADKCKVYLLGAHSL 60
DB 26 ILGGEAEAHAPRYMASVOLGNHLCGGVAVABOMVLASNCLEDADKCKVYLLGAHSL 85
QY 61 SQEPKSLRYDLVARVPDSQPTDIDHLLGLQSEKATLGPVRLPMQVRDVRAP 120
DB 86 SQEPKSLRYDLVARVPDSQPTDIDHLLGLQSEKATLGPVRLPMQVRDVRAP 145
QY 121 TLGVAGVGVYNHAGRRPDSQVHLVLPVLDRTACNRTHDGAITPRLCMASNRROSK 180
DB 146 TLGVAGVGVYNHAGRRPDSQVHLVLPVLDRTACNRTHDGAITPRLCMASNRROSK 205
QY 181 GDSGCPVLCGGVLEGVTSQSRVCGNRKKKPGITRVASVAAWTDVSLA 228
DB 206 GDSGCPVLCGGVLEGVTSQSRVCGNRKKKPGITRVASVAAWTDVSLA 253
RESULT 2.
ID CPAD_PIG STANDARD; PRT: 259 AA.
DC P51779;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement factor D precursor (EC 3.4.21.46) (C3 convertase activator)
DE Properdin factor D (Adipsin).
OS Homo sapiens
OS scrofa (Pig).
OS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;
[1]
RN SEQUENCE FROM N.A.
RC TISSU=Adipose tissue;
RC NCBI_TaxID=9823;
RN Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN SEQUENCE OF 69-259 FROM N.A.
RC TISSU=Adipose tissue;
RA Nicolas N.;
RA Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: FACTOR D (FACD) IS A MEMBER OF THE LATEIN WHICH IS COMPLEXED
CC WITH THE C3 CONVERTASE OF THE ALTERNATE PATHWAY. ITS FUNCTION IS
CC TO CLEAVE FACTOR B INTO ACTIN AND BETA-2-MICROGLOBULIN. IT BECOMES
CC HOMOLOGOUS TO THAT OF CLS IN THE CLASSICAL PATHWAY (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Cleaves component factor B (Arg-1-Lys) when in
CC complex with C3b or with cobra venom factor (CVF).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
-----
This SwissProt entry is copyrighted. It is produced through a collaboration
between the Swiss Institute for Bioinformatics and the EMBL outstations
of the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announcement/
 CC on send an email to license@isb-sib.ch.)

CC ENBL: U2948; AAT3627.1; -;
 CC HSP: P40508; CA088844.1; -;
 CC HSP: P40746; IDSU.
 CC MRPS: S01191; -;
 CC InterPro: IPR001314; Chymotrypsin.
 CC Pfam: PF00089; trypsin; 1.
 CC PRINTS: PR00722; CHYMOTRYPSIN.
 CC SMART: SM00020; Tryp.SPC; 1.
 CC PROSITE: PS00240; TRYPSIN.DOM; 1.
 CC PROSITE: PS00134; TRYPSIN.HIS; 1.
 CC PROSITE: PS00135; TRYPSIN.SER; 1.
 CC Complement alternate pathway: Plasma; Hydrolyase; Serine protease;
 CC Zymogen: Glycoprotein; Signal.
 CC SIGNAL: 1 21 POTENTIAL.
 CC CHAIN 22 26 ACTIVATION PEPTIDE (POTENTIAL).
 CC PROPEP 27 259 COMPLEMENT FACTOR D.
 CC ACT_SITE 67 167 CHARGE RELAY SYSTEM.
 CC ACT_SITE 165 215 CHARGE RELAY SYSTEM.
 CC FT DISULFID 203 268 BY SIMILARITY.
 CC FT DISULFID 149 215 BY SIMILARITY.
 CC FT DISULFID 180 196 BY SIMILARITY.
 CC FT DISULFID 205 230 BY SIMILARITY.
 CC FT DISULFID 205 230 BY SIMILARITY.
 CC SEQUENCE 259 AA: 27763 MW: 0121AAE0E34CAIED CRC64:
 Query Match 87.24; Score 1056; DB 1; Length 259;
 Best Local Similarity 84.24; Pred. No. 99-93; Mismatches 18; Indels 0; Gaps 0;
 Matches 192; Conservative 18; Mismatches 18; Indels 0; Gaps 0;
 QY 1 ILGGRERARHPYMASVYOLNAGHLCGGVLYAQQWYLSAAHCLDEADQKQVLLGAHSL 60
 Db 27 ILGGRERARHPYMASVYOLNAGHLCGGVLYAQQWYLSAAHCLDEADQKQVLLGAHSL 86
 QY 61 SQPEPSKRLYDVLRAVPHPSQPTIDHLLQLLSEKATLGAVRPLQVREVPAG 120
 Db 87 SQPEPSKRLYDVLRAVPHPSQPTIDHLLQLLSEKATLGAVRPLQVREVPAG 146
 QY 121 TLCDVAGGVNHAGRRPDSLOHVLPLVDLRATCNRRTHGATITRLMACSNRRDCK 180
 Db 147 TLCDVAGGVNHAGRRPDSLOHVLPLVDLRATCNRRTHGATITRLMACSNRRDCK 206
 QY 181 GDSGGLPLVCGGVGGVYVTSGRVCGNKKPGIVYTRVASTAAVDSVLA 228
 Db 207 GDSGGLPLVCGGVGGVYVTSGRVCGNKKPGIVYTRVASTAAVDSVLA 254

RESULT 3
 PFAD: 01-JUL-1993 (Rel. 26, Created)
 AC P12038; (Rel. 35, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 35, Last sequence update)
 DE Complement factor D precursor (EC 3.4.21.46) (C3 convertase activator)
 DE (Properdin factor D) (Adipsin) (Endogenous vascular elastase)
 GN DR ADN. (Rel. 26, Created)
 CC Eukaryotic; Mammalian; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID:10116;
 RN SEQUENCE FROM N.A.
 RX MIMDMEP-94365184; PubMed-8083356;
 RX Zhu L., Wigle D., Hinek A., Kobayashi J., Ye C., Zuker M.,
 RX 1997. The endogenous vascular elastase (Adipsin) is a novel enzyme related to the serine proteinase adipsin.
 RT "progression of monocrotaline-induced pulmonary hypertension in rats
 RT is a novel enzyme related to the serine proteinase adipsin.";

J. Clin. Invest. 94:1163-1171(1994).
 RN [2]
 AP SEQUENCE OF 26-55;
 BA HEADLINE-92061993; PubMed-1953671;
 BL ENBL: U2948; AAT3627.1; -;
 BL HSP: P40508; CA088844.1; -;
 BL HSP: P40746; IDSU.
 BL MRPS: S01191; -;
 BL InterPro: IPR001314; Chymotrypsin.
 BL Pfam: PF00089; trypsin; 1.
 BL PRINTS: PR00722; CHYMOTRYPSIN.
 BL SMART: SM00020; Tryp.SPC; 1.
 BL PROSITE: PS00240; TRYPSIN.DOM; 1.
 BL PROSITE: PS00134; TRYPSIN.HIS; 1.
 BL PROSITE: PS00135; TRYPSIN.SER; 1.
 BL Complement alternate pathway: Plasma; Hydrolyase; Serine protease;
 BL Zymogen: Glycoprotein; Signal.
 BL SIGNAL: 1 20 POTENTIAL.
 BL CHAIN 21 23 ACTIVATION PEPTIDE.
 BL PROPEP 22 259 COMPLEMENT FACTOR D.
 BL ACT_SITE 67 167 CHARGE RELAY SYSTEM.
 BL ACT_SITE 165 215 CHARGE RELAY SYSTEM.
 BL FT DISULFID 203 268 BY SIMILARITY.
 BL FT DISULFID 149 215 BY SIMILARITY.
 BL FT DISULFID 180 196 BY SIMILARITY.
 BL FT DISULFID 205 230 BY SIMILARITY.
 BL FT DISULFID 205 230 BY SIMILARITY.
 BL SEQUENCE 263 AA: 28442 MW: D57131356A13CE0 CRC64:
 Query Match 71.84; Score 869.5; DB 1; Length 263;
 Best Local Similarity 69.04; Pred. No. 4.2e-75;
 Matches 158; Conservative 31; Mismatches 39; Indels 1; Gaps 1;
 QY 1 ILGGRERARHPYMASVYOLNAGHLCGGVLYAQQWYLSAAHCLDEADQKQVLLGAHSL 59
 Db 26 ILGGRERARHPYMASVYOLNAGHLCGGVLYAQQWYLSAAHCLDEADQKQVLLGAHSL 85
 QY 60 SQPEPSKRLYDVLRAVPHPSQPTIDHLLQLLSEKATLGAVRPLQVREVPAG 119
 Db 86 SQPEPSKRLYDVLRAVPHPSQPTIDHLLQLLSEKATLGAVRPLQVREVPAG 145
 QY 120 TLCDVAGGVNHAGRRPDSLOHVLPLVDLRATCNRRTHGATITRLMACSNRRDCK 179
 Db 146 TLCDVAGGVNHAGRRPDSLOHVLPLVDLRATCNRRTHGATITRLMACSNRRDCK 205
 QY 180 KDSGGLPLVCGGVGGVYVTSGRVCGNKKPGIVYTRVASTAAVDSVLA 228
 Db 206 KDSGGLPLVCGGVGGVYVTSGRVCGNKKPGIVYTRVASTAAVDSVLA 254

RESULT 4

CEAD_MOUSE STANDARD; PRT; 259 AA.
 AC P01953.061280;
 DT 23-OCT-1986 (Rel. 02, Created)
 DT 23-OCT-1986 (Rel. 02, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Complement factor D precursor (EC 3.4.21.46) (C3 convertase activator)
 OS (Properdin factor D) (Adipsin) (28 kDa protein, adipocyte).
 DE Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RX PHILIPINE-9627684; PubMed=3015943;
 RA PHILIPINE-9627684;
 RA "The nucleotide sequence of three genes participating in the adipose
 differentiation of 3T3 cells.";
 RT J. Biol. Chem. 261:10821-10827(1986).
 F [2]
 F SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RA PHILIPINE-9627684; PubMed=3024123;
 RA Min H. H. 1996; Adipocyte serine protease:
 RA "Adipsin", the adipocyte serine protease: gene structure and control
 of expression by tumor necrosis factor.";
 RT Nucleic Acids Res. 14:8879-8892(1986).
 RM [3]
 RM SEQUENCE FROM N.A.
 RP Cook K.S.; Groves D.L.; Min H.Y.; Spiegelman B.M.;
 RX MEDLINE=46016726; PubMed=3901003;
 RA "A developmentally regulated mRNA from 3T3 adipocytes encodes a novel
 serine protease homologue.";
 RT Proc. Natl. Acad. Sci. U.S.A. 82:6480-6484(1985).
 CC -1- FUNCTION: FACTOR D CLEAVES FACTOR B WHEN THE LATTER IS COMPLEXED
 WITH FACTOR C3B. ACTIVATING THE C3BB COMPLEX, WHICH THEN BECOMES
 THE C3 CONVERTASE. THIS COMPLEX THEN CLEAVES FACTOR B INTO TWO
 HOMOLOGOUS TO THAT OF C1S IN THE CLASSICAL PATHWAY.
 CC -1- CATALYTIC ACTIVITY: Cleaves component factor B (Arg-1-Lys) when in
 complex with C3b or with cobra venom factor (CVF).
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 produced by alternative splicing.
 CC -1- PTM: N-GLYCOSYLATED.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-stb.ch/announce/>
 or <http://www.ebi.ac.uk/seqdb/doc/seqdb.html>).
 CC -----
 CC ENBL; M1768; AAA0486.1;
 CC ENBL; X04623; CAA23878.1;
 CC ENBL; X04673; CAA28379.1;
 CC ENBL; M13386; AAA37262.1;
 CC PIR; C25952; WMS28.
 CC PIR; F00416; I1850.
 CC MEROPS; S01.19;
 CC MGD; MGI:87931; Adn.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; TRYP_SPE; 1-0.
 CC PROSITE; PS00134; TRYPSPN; 1.
 CC PROSITE; PS00135; TRYPSPN_SER; 1.
 CC Complement alternate pathway: Plasma; Hydrolase; Serine protease;
 KW Zymogen; Glycoprotein; Alternative splicing; Signal.
 CC SIGNAL 1 20 OR 19 (POTENTIAL).

FT PROPEP 21 25 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 26 259 COMPLEMENT FACTOR D.
 FT ACT_SITE 196 196 CHANGE RELAY SYSTEM.
 FT ACT_SITE 197 197 CHANGE RELAY SYSTEM.
 FT ACT_SITE 209 209 CHANGE RELAY SYSTEM.
 FT DISULFID 51 67 BY SIMILARITY.
 FT DISULFID 149 215 BY SIMILARITY.
 FT DISULFID 180 196 BY SIMILARITY.
 FT DISULFID 205 230 BY SIMILARITY.
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 20 20 MISSING (IN ISOFORM 2).
 SO SEQUENCE 259 AA; 28057 MW; 8C3A952561247DF9 CRC64;
 Query Match. Score 844.5; DB 1; Length 259;
 Best Local Similarity 66.58; Pct. No. 9.5e-73;
 Matches 151; Conservative 33; Mismatches 42; Indels 1; Gaps 1;
 Oy 1 IIGREAEAHARPYASVQNGAHLCGGVLVAEQWLSAAHCLDEAA-DGKVVVLGHAHS 59
 Dy 26 IIGREAEAHARPYASVQNGAHLCGGVLVAEQWLSAAHCLDEAA-DGKVVVLGHAHS 85
 Oy 60 LSQSPGKRLYDVRVAVPDSQPTDIDHLLQSLKATGLPAPVPLPQRPVQVDPVAP 119
 Dy 86 LSAPKPYKRYDQSVQVWPRDPSLEDDLLFLKLSQNASLGPVRPLPQYKDEVEP 145
 Oy 120 GTCLDVAGNGVNHAGRRRPSLQHVLFVLDVLRATCNRTPHDGAITERLACASNERDSC 179
 Dy 146 GTCLDVAGNGVNHAGRRRPSLQHVLFVLDVLRATCNRTPHDGAITERLACASNERDSC 205
 Oy 180 KDSGSGPLVCGVLEGVVSGSRVCGNKKPGITRVASYAAMDSV 226
 Dy 206 RDSGSGPLVCGDAVGVTGWSRVCNGKPGVTRVSSYRWMIENI 252
 RESULT 5
 ID GRAM_HUMAN STANDARD; PRT; 257 AA.
 AC P51124;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Granzyme M precursor (EC 3.4.21.-) (Met-ase) (Natural killer cell
 cytotoxic granule-associated serine protease).
 GN GZMM OR VET1
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RX SEQUENCE FROM N.A.
 RX MEDLINE=94065173; PubMed=8245461;
 RA Smyth M.J.; Savers T.J.; Willtrout T.; Powers J.C.; Trapani J.A.;
 RA "Met-ase: cloning and distinct chromosomal location of a serine
 protease preferentially expressed in human natural killer cells.";
 RT J. Immunol. 151:6195-6205(1993).
 RL [2]
 RL SEQUENCE FROM N.A.
 RX MEDLINE=94065173; PubMed=7711495;
 RA Pilat D.; Flak T.M.; Obermair S.; Schrobek B.; Zimmer M.; Wekerle H.;
 RA Lichter P.; Jenne D.E.;
 RA "The human Met-ase gene (GZMM): Structure, sequence, and close
 physical linkage to the serine protease gene cluster on 19p13.3.";
 RT Genomics 24:445-450(1994).
 RL [3]
 RL SEQUENCE FROM N.A.
 RP Smyth M.J.; (JAN-1994) to the EMBL/Genbank/DBJ databases
 RX Submission: CLEAVES PEPTIDE SUBSTRATES AFTER METHIONINE, LEUCINE,
 CC -1- FUNCTION: CLEAVES PEPTIDE SUBSTRATES AFTER METHIONINE, LEUCINE,
 CC AND NORLEUCINE.

Result	Query No.	Score	Match	Query Length	DB ID	Description
1	1211	100.0	253	4	08W2B4	Q8W2B4 homo sapien
2	554	45.7	255	13	Q34289	Q34289 salvelinus
3	457.5	37.8	256	11	Q9R0K0	Q9R0K0 mus musculus
4	436	36.0	264	11	Q08643	Q08643 mus musculus
5	437.5	35.7	240	13	Q98RH0	Q98RH0 engraulis
6	429	35.4	244	13	Q42159	Q42159 petromyzon
7	429	35.4	245	13	Q42160	Q42160 petromyzon
8	428	35.3	251	6	09JG22	09JG22 bos taurus
9	425.8	32.1	251	11	Q88954	Q88954 rattus norv
10	421.5	31.8	256	11	Q88954	Q88954 mus musculus
11	421.5	34.8	253	11	Q91782	Q91782 mus musculus
12	419	34.6	247	11	Q9CPW7	Q9CPW7 mus musculus
13	418	34.5	247	13	Q42608	Q42608 petromyzon
14	417	34.4	247	13	Q42158	Q42158 petromyzon
15	414.5	34.2	241	13	Q98RT9	Q98RT9 engraulis
16	408.5	33.7	255	4	Q96R00	Q96R00 homo sapien

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 22:08:31 : Search time 46 Seconds
(without alignments)
427,682 Million cell updates/sec

Title: US-09-821-255-2

Perfect score: 1211

Sequence: 1 ILGREAERAHPRVMSYQL.....KFGIYTRVASYAAMDVSLA 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs. 86286685 residues

il number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1211	100.0	228	10	US-09-821-255-2
2	1056	87.2	233	10	US-09-821-255-4
3	456.5	37.7	264	9	US-10-114-893-26
4	451	37.2	284	10	US-09-888-615-112
5	447	36.9	283	9	US-10-174-590-272
6	447	36.9	283	9	US-10-176-758-272
7	447	36.9	283	9	US-10-175-737-272
8	447	36.9	283	9	US-10-173-716-272
9	447	36.9	283	9	US-10-173-752-272
10	447	36.9	283	9	US-10-176-482-272
11	447	36.9	283	9	US-10-176-757-272
12	447	36.9	283	9	US-10-176-913-272
13	447	36.9	283	9	US-10-180-552-272
14	447	36.9	283	9	US-10-180-557-272
15	447	36.9	283	9	US-10-173-750-272
16	447	36.9	283	9	US-10-174-579-272
17	447	36.9	283	9	US-10-174-579-272
18	447	36.9	283	9	US-10-174-579-272
19	447	36.9	283	9	US-10-174-582-272

20	447	36.9	283	9	US-10-174-588-272	Sequence 272, App
21	447	36.9	283	9	US-10-175-739-272	Sequence 272, App
22	447	36.9	283	9	US-10-175-740-272	Sequence 272, App
23	447	36.9	283	9	US-10-175-743-272	Sequence 272, App
24	447	36.9	283	9	US-10-176-482-272	Sequence 272, App
25	447	36.9	283	9	US-10-176-482-272	Sequence 272, App
26	447	36.9	283	9	US-10-176-747-272	Sequence 272, App
27	447	36.9	283	9	US-10-176-750-272	Sequence 272, App
28	447	36.9	283	9	US-10-176-985-272	Sequence 272, App
29	447	36.9	283	9	US-10-176-987-272	Sequence 272, App
30	447	36.9	283	9	US-10-176-991-272	Sequence 272, App
31	447	36.9	283	9	US-10-176-993-272	Sequence 272, App
32	447	36.9	283	9	US-10-176-993-272	Sequence 272, App
33	447	36.9	283	9	US-10-184-658-272	Sequence 272, App
34	447	36.9	283	9	US-10-173-695-272	Sequence 272, App
35	447	36.9	283	9	US-10-173-697-272	Sequence 272, App
36	447	36.9	283	9	US-10-173-705-272	Sequence 272, App
37	447	36.9	283	9	US-10-174-576-272	Sequence 272, App
38	447	36.9	283	9	US-10-174-582-272	Sequence 272, App
39	447	36.9	283	9	US-10-174-582-272	Sequence 272, App
40	447	36.9	283	9	US-10-175-747-272	Sequence 272, App
41	447	36.9	283	9	US-10-176-481-272	Sequence 272, App
42	447	36.9	283	9	US-10-176-485-272	Sequence 272, App
43	447	36.9	283	9	US-10-176-487-272	Sequence 272, App
44	447	36.9	283	9	US-10-176-493-272	Sequence 272, App
45	447	36.9	283	9	US-10-176-756-272	Sequence 272, App

ALIGNMENTS

```

RESULT 1
: Seq ID 1 255-2
: Sequence 2 Application US/09821255
: Patent No. US20020081293A1
: GENERAL INFORMATION:
: APPLICANT: Michael S.C. Fung
: APPLICANT: Cecily R.Y. Sun
: APPLICANT: INVENTOR: Cecily R.Y. Sun
: FILE REFERENCE: 98-22
: INHIBITORS OF Complement Activation
: CURRENT APPLICATION NUMBER: US/09/821.255
: PRIOR FILING DATE: 2001-03-29
: PRIOR APPLICATION NUMBER: 60/075.328
: PRIOR FILING DATE: 1998-02-20
: PRIOR APPLICATION NUMBER: 09/253.689
: FILING DATE: 1999-02-20
: NUMBER OF SEQ ID NO. 1
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 228
: TYPE: PRT
: ORGANISM: human
US-09-821-255-2

```

Query Match 100.0% Score 1211; DB 10; Length 228;
Best Local Similarity 100.0%; From No. 76-106;
Matches 228; Conservative 0; Mismatches 0; Gaps 0;

QY	1	ILGREAERAHPRVMSYQLNHLGCGVIVABOWVLSAAHCLDADAGKGVLLGAHSL	60
DB	1	ILGREAERAHPRVMSYQLNHLGCGVIVABOWVLSAAHCLDADAGKGVLLGAHSL	60
QY	61	SOPEPSKRLTDVLAHPHDSQDPTIDHLLLOLSEKATLGPVPLPQKRDVAVP	120
DB	61	SOPEPSKRLTDVLAHPHDSQDPTIDHLLLOLSEKATLGPVPLPQKRDVAVP	120
QY	121	TLCDVAGWGIVNHAGRRPDSLOHVLPLDPTCNRRTHDCAITERLAKCSNRDCK	180
DB	121	TLCDVAGWGIVNHAGRRPDSLOHVLPLDPTCNRRTHDCAITERLAKCSNRDCK	180
QY	181	CDSCGGLVCGVLEGVVTSQSGVCGNRKKPGVITRWASYAAMDVSLA 228	

Db 181 GDSGGLVCGVLEGVVTSGRVCGNRKKPGIYTRVASYAAMIDSVLA 228

RESULT 2

US-09-821-255-4

; Sequence 4, Application US/09821255

; Patent No. US20020081293A1

; GENERAL INFORMATION: Inventor: C. Fung

; APPLICANT: MALLIN, R.C. Sun

; APPLICANT: BELL, R.C. Sun

; APPLICANT: Cecily R.Y. Sun

; TITLE OF INVENTION: Inhibitors of Complement Activation

; FILE REFERENCE: 98-2A

; CURRENT APPLICATION NUMBER: US/09/821,255

; CURRENT FILING DATE: 2001-03-29

; PRIORITY FILING DATE: 1998-07-07/5, 328

; PRIOR FILING DATE: 1998-07-07/5, 328

; PRIOR APPLICATION NUMBER: 09/253,689

; PRIOR FILING DATE: 1999-02-20

; NUMBER OF SEQ ID NOS: 15

; IFTNAME: FastSeq for Windows Version 4.0

; SEQ ID NO: 233

; TYPE: PRT

; ORGANISM: pig

US-09-821-255-4

Query Match 87.24; Score 1056; DB 10; Length 233;

Best Local Similarity 84.24; Pred. No. 1.7e-91;

Matches 192; Conservative 18; Mismatches 18; Indels 0; Gaps 0;

QY 1 ILGREGAENARHPYASVQNLGALCGVLAQVWLSAHCLEDAADCKVQVLLGAHSI 60

Db 1 ILGQGEAKSHERPYSVQVNGKHVCGGELVNSEQWLSAHCLEDAVAKIKVLLGAHSI 60

QY 61 SPPESKRLYDLRAVPHPSQDPTIDHLLQLLSEKATLGPVRLPQWRDVRDAPG 120

Db 61 SPPESKRLYDLRAVPHPSQDPTIDHLLQLLSEKATLGPVRLPQWRDVRDAPG 120

QY 121 TLCDVAGMGVNHAGRPDSQHLVLPVLDRTCNRRTHDGAITERLMAESNRDSCG 180

Db 121 TLCDVAGMGVYHGRPDRLQHLVLPVLDRTCNRTYHDGTITERLMAESNRDSCG 180

QY 181 GDSGGLVCGVLEGVVTSGRVCGNRKKPGIYTRVASYAAMIDSVLA 228

Db 181 GDSGGLVCGVLEGVVTSGRVCGNRKKPGIYTRVASYAAMIDSVLA 228

R "T 3

; -114-993-26

; Sequence 26, Application US/10114893

; Patent No. US20020193567A1

; GENERAL INFORMATION: Inventor: John Kenneth

; APPLICANT: Jacoby, John M.

; APPLICANT: McCoy, John M.

; APPLICANT: Lavallie, Edward R.

; APPLICANT: Collins-Racie, Lisa A.

; APPLICANT: Evans, Cheryl

; APPLICANT: Marberg, David

; APPLICANT: Freacy, Marilee R.

; APPLICANT: Spaulding, Vikki

; APPLICANT: Carlin-Duckett, McKeough

; APPLICANT: Genetics Institute, Inc.

; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

; FILE REFERENCE: G1 6000-10A

; CURRENT APPLICATION NUMBER: US/10/114,893

; CURRENT FILING DATE: 1999-10-06

; EARLIER APPLICATION NUMBER: 09/413,232

; EARLIER FILING DATE: 1999-10-06

; NUMBER OF SEQ ID NOS: 321

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 26

; LENGTH: 264

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-114-893-26

Query Match

Best Local Similarity 37.74; Score 456.5; DB 9; Length 264;

Matches 100; Conservative 40; Mismatches 85; Indels 9; Gaps 7;

QY 1 ILGREGAENARHPYASVQNLGALCGVLAQVWLSAHCLEDAADCKVQVLLGAHSI 59

Db 27 ILGGEVSHSRFHWASIQVGHVCGSVLLDQWVLTAAIKQVPTKQSPVVLGAHSI 86

QY 60 LQSPESKRLYDLRAVPHPSQDPTIDHLLQLLSEKATLGPVRLPQWRDVRDAPG 119

Db 87 LSKNEASQKTLERKIPFSPVTSQPSQNDIMLVKLOTAANKLKHVRL-IRSKTSLS 145

QY 120 GTLLCDVAGMGVNHAGRPDSQHLVLPVLDRTCNRRTHDGAITERLMAESNRDSCG 174

Db 146 GTFKCVTGKATDPDSLRPSOTLRVTVVLSKRLCNSQSYNGDPFIKDYKACGAKG 205

QY 175 RDSCGSGSGPLVCGVLEGVVTSGRVCGNRKKPGIYTRVASYAAMIDSVLA 227

Db 206 QKDSCKGSGPLVCGVLEGVVTSGRVCGNRKKPGIYTRVASYAAMIDSVLA 258

RESULT 4

US-09-888-615-112

; Sequence 112, Application US/0988615

; Patent No. US20020064856A1

; GENERAL INFORMATION: Inventor: GREGORY

; APPLICANT: PLOMAN, GREGORY

; APPLICANT: BRYTE, DAVID

; APPLICANT: CAENEPEEL, SEAN

; APPLICANT: CHARYDCZAK, GLEN

; APPLICANT: MANNING, GERARD

; APPLICANT: SUDANSANNA, SUCHA

; APPLICANT: CHOWDHURY, S. PROTEASES

; FILE REFERENCE: 03/9602/121

; CURRENT APPLICATION NUMBER: US/09/888,615

; CURRENT FILING DATE: 2001-06-26

; PRIOR FILING DATE: 2000-06-26

; NUMBER OF SEQ ID NOS: 150

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO: 244

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-888-615-112

Query Match

Best Local Similarity 37.28; Score 451; DB 10; Length 284;

Matches 100; Conservative 34; Mismatches 86; Indels 12; Gaps 6;

QY 1 ILGREGAENARHPYASVQNLGALCGVLAQVWLSAHCLEDAADCKVQVLLGAHSI 58

Db 35 ILGGEVSHSRFHWASIQVGHVCGSVLLDQWVLTAAIKQVPTKQSPVVLGAHSI 92

QY 59 SLPSPESKRLYDLRAVPHPSQDPTIDHLLQLLSEKATLGPVRLPQWRDVRDAPG 116

Db 93 VLSSTAEPFQGVFGDITLTHDVPWTHANDICLLDGLGNSVLPVGLLRFGARR-AP 151

QY 117 VAPGTLCDVAGMGVNHAGRPDSQHLVLPVLDRTCNRRTHDGAITERLMAESNRDSCG 173

Db 152 PTAETRCVARGVGFDEFELPGLMEAKRVLPDQVGN--SSNKGHLTLMCTRSGDS 209

QY 174 NRDSCKGSGSGPLVCGVLEGVVTSGRVCGNRKKPGIYTRVASYAAMIDSVLA 227

Db 210 HRGFTCSADSGPLVCGVLEGVVTSGRVCGNRKKPGIYTRVASYAAMIDSVLA 263